序 言

原来准备写一部《高等数学引论》，共六、七卷．其中第一卷第一、二分册已于 1963 年问世．但由于十年浩劫，其它手稿大部分遭到一"抄"，二"盗"，三"失散"的命运．现在检查劫余，已所剩无几了．于 1981 年出版了第二卷第一分册．本拟鼓其余勇完成原来计划，但实事求是地估计之后，看来所散失的手稿归来无望，重新补写完成全书，诚恐是时不我待，力不从心的愿望了．无已，作两步打算，先把 1962 年在中国科技大学讲授过的，铅印而未逸散的部分出版，以后，再就力所能及进行写作，陆续出版(在同一或不同书名下)．

好在这一部分是有它的独立性的，讲授时的客观情况是：既要照顾到初学同学的水平，又要使前者不越级，后者不觉得是简单重复，总的原则是重用矩阵，实现"1，2，3；n；∞"讲授法的第二步．即，准备前几卷是讲一、二、三个变数，一、二、三维空间；而这一卷是讲 n 个变数或 n 维空间．本卷原稿缺第十、十一、十二，三章，据回忆这三章是讲 n 维空间微分几何学的，原稿虽失，但读者不妨以第一卷空间曲线的微分几何为模型，运用正交群下斜对称方阵的分类而获得 n 维空间曲线的微分性质．这是一个好习题，如果能做得出，则可把正交群改为其它群，而研究其微分不变性质．

岁月无多，不得不计日图效，错谬之处请读者指正．

华罗庚

1981 年 10月11日

又 序

我非常感谢科学出版社能够出版这残余的手稿．这是对科学工作的珍视．但编辑就为此增加了不少麻烦，花费了不少精力，我在此致谢．

当年，在科技大学编写此书时，龚昇同志给我许多帮助，在寻找遗失稿件时，他还多方尽力．

对龚昇同志和负责校对的裴定一同志，我在此敬致谢忱．

华罗庚

1983年 9 月 9 月

目　　录

第一章 线性方程组与行列式(复习提纲)

§1. 线性方程组

考虑齐次方程组:

$$\sum_{j=1}^{n} a_{ij}x_j = 0, \quad i = 1, \cdots, n, \tag{1}$$

这儿 a_{ij} 是复数(或实数) $x_1, \cdots, x_n$ 是未知数. 方程组(1)显然有一个解

$$x_1 = \cdots = x_n = 0. \tag{2}$$

这个解称为显见解.

研究齐次方程组的基本问题是: 除显见解外,(1)是否还有其他解? 能否定出所有的解来?

非齐次方程组

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \cdots, n \tag{3}$$

的基本问题是: (3) 是否有解? 能否定出所有的解来.

如果 (3) 有一个解 $(x_1^0, \cdots, x_n^0)$, 即

$$\sum_{j=1}^{n} a_{ij}x_j^0 = b_i,$$

则

$$\sum_{j=1}^{n} a_{ij}(x_j - x_j^0) = 0.$$

命 $y_j = x_j - x_j^0$, 则 y_j 是 (1) 的解. 所以解非齐次方程组的问题一变而为两个: 首先是是否有解,其次定出齐次方程组 (1) 的所有的解来.

关于是否有解有次之重要结果:

如果(1)有非显见解,则(3)不能对所有的 $b_1, \cdots, b_n$ 都有解.

如果(1)仅有显见解,则(3)对任意的 $b_1, \cdots, b_n$ 都有解.

§2. 消去法

解线性方程组(3)的方法我们着重复习一下 Gauss 消去的原则. 以四个未知数、四个方程为例

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 = a_{15}, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 = a_{25}, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 = a_{35}, \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 = a_{45}. \end{cases} \tag{1}$$

将(1)中的第一个方程除以系数 a_{11} (它叫做"主导"元素)并令

$$b_{1j} = \frac{a_{1j}}{a_{11}} \quad (j > 1), \tag{2}$$

则得一个新的方程

$$x_1 + b_{12}x_2 + b_{13}x_3 + b_{14}x_4 = b_{15}. \tag{3}$$

再由方程(3)及(1)中的后面三个方程消去 x_1，这样便得到了一个辅助方程组，它包括具有三个未知数的三个方程，此种消去法易于施行，只须顺次将方程(3)乘以 a_{21}，a_{31}，a_{41} (也就是乘以第二、第三和第四行的"主导"元素)，再由(1)中的对应方程减去此式即可，消去一个未知数以后所得的新方程组，其系数用 $a_{ij,1}$ 代表：

$$a_{ij,1} = a_{ij} - a_{i1}b_{1j}(i, j \geqslant 2). \tag{4}$$

其次将新方程组中的第一式除以它的"主导"元素 $a_{22,1}$，则得方程

$$x_2 + b_{23,1}x_3 + b_{24,1}x_4 = b_{25,1}. \tag{5}$$

其中

$$b_{2j,1} = \frac{a_{2j,1}}{a_{22,1}} \quad (j > 2), \tag{6}$$

然后仿照前面的方法继续进行，我们便得到了一组具有两个未知数的两个方程，它们的系数呈如次的形式：

$$a_{ij,2} = a_{ij,1} - a_{i2,1}b_{2j,1} \quad (i, j \geqslant 3). \tag{7}$$

将这组方程的第一式除以主导元素 $a_{33,2}$，并令

$$b_{3j,2} = \frac{a_{3j,2}}{a_{33,2}} \quad (j > 3), \tag{8}$$

则得方程

$$x_3 + b_{34,2}x_4 = b_{35,2}.$$

最后再做一步，即可得出一个方程，它只含一个未知数，而其系数为 $a_{44,3}$，将这个方程除以 $a_{44,3}$，则得

$$x_4 = b_{45,3}.$$

将具有系数 $b_{ij,i-1}$ $(j > i)$ 的一切方程合并，便得到一个三角形的方程组，它与原有的方程组等价；它的解就是原有方程组的解，我们要注意，上述方法只有当所有的"主导"元素都不等于零时才能使用.

我们把求三角形方程组的系数的手续称为正面过程，而把求三角形方程组的解的手续称为反面过程(参看附表).

我们还要讲一下验算的方法，用代换 $\bar{x}_i = x_i + 1$，则我们得到一组以 $\bar{x}_i$ 为变数的方程组，它的系数与原来的方程相同，而它的常数项等于原方程的系数与常数项之和，我们可以同时计算这两个方程组，求出解 $\bar{x}_i$，并视其是否等于 $x_i + 1$，这就是验算方法.

现在简单地说明一下附表：

正面过程是用如下方法来施行的，写出矩阵系数(包括常数项与核验和)将第一行除以主导元素，并将结果写成矩阵最末一行，再求出第一个辅助系数 $a_{ij,1}(i, j \geqslant 2)$：从已知矩阵任取一个元素，由它减去一个乘积——就是上述元素所在的那一行的主导元素与上述元素所在的那一列的最末元素的乘积，重复施行这种手续，当我们得出了仅含一行的矩阵时，正面过程便完成了.

附 表 1

x_1	x_2	x_3	x_4		Σ						Σ
a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	1.00	0.42	0.54	0.66	0.3	2.92
a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	0.42	1.00	0.32	0.44	0.5	2.68
a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	a_{36}	0.54	0.32	1.00	0.22	0.7	2.78
a_{41}	a_{42}	a_{43}	a_{44}	a_{45}	a_{46}	0.66	0.44	0.22	1.00	0.9	3.22
1	b_{12}	b_{13}	b_{14}	b_{15}	b_{16}	1	0.42	0.54	0.66	0.3	2.92
	$a_{22,1}$	$a_{23,1}$	$a_{24,1}$	$a_{25,1}$	$a_{26,1}$		0.82360	0.09320	0.16280	0.37400	1.45360
	$a_{32,1}$	$a_{33,1}$	$a_{34,1}$	$a_{35,1}$	$a_{36,1}$		0.09320	0.70840	−0.13640	0.53800	1.20320
	$a_{42,1}$	$a_{43,1}$	$a_{44,1}$	$a_{45,1}$	$a_{46,1}$		0.16280	−0.13640	0.56440	0.70200	1.29280
	1	$b_{23,1}$	$b_{24,1}$	$b_{25,1}$	$b_{26,1}$		1	0.11316	0.19767	0.45410	1.76493
		$a_{33,2}$	$a_{34,2}$	$a_{35,2}$	$a_{36,2}$			0.69785	−0.15482	0.49568	1.03871
		$a_{43,2}$	$a_{44,2}$	$a_{45,2}$	$a_{46,2}$			−0.15482	0.53222	0.62807	1.00547
		1	$b_{34,2}$	$b_{35,2}$	$b_{36,2}$			1	−0.22185	0.71030	1.48844
			$a_{44,3}$	$a_{45,3}$	$a_{46,3}$				0.49787	0.73804	1.23591
			1	x_4	$\bar{x}_4$				1	1.48240	2.48240
		1		x_3	$\bar{x}_3$			1		1.03917	2.03916
	1			x_2	$\bar{x}_2$		1			0.04348	1.04348
1				x_1	$\bar{x}_1$	1				−1.25780	−0.25779

在反面过程中，我们利用包含 1 的各行而由最末一行开始，精确地说，在这些行的最后一行里，我们从常数项的一列中得到了最后一个未知量的值，而在核验列中得到了核验值，然后可以逐次得出各个未知量的值，只要由倒数第二列的元素减去对应系数 b 与前面所得未知量 i 值的乘积即可，在表格的末尾写出 1 字，可以帮助我们找出在所要各行中对应于已知 x 的系数，例如

$$x_2 = b_{25,1} - b_{23,1}x_3 - b_{24,1}x_4$$
$$= 0.45410 - 0.11316 \times 1.03917 - 0.19767 \times 1.48240 = 0.04348.$$

最后，我们还要指出用这种方法解 n 个变数的线性方程组所需的乘法与除法的运算次数为 $\frac{n}{3}(n^2 + 3n - 1)$.

§ 3. 消去法的几何解释

先看两个变数的情况.

$$l: \quad ax + by = c, \qquad l': \quad a'x + b'y = c',$$

在平面上各表示一条直线，两个直线有一交点：消去 y，得出仅有 x 的方程，这是这个交点在 x 轴上的投影.

也可以这样看：第一、二方程各表示一条直线 l 与 l'. 由方程

$$\lambda(ax + by - c) + \mu(a'x + b'y - c') = 0$$

定义出一族直线，这些直线由 $\lambda l+\mu l'$ 表示，这些直线有一个重要性质，就是通过 l 与 l' 的交点，不难证明：反过来，凡是通过 l 与 l' 的交点的直线也在这族之中，在这族直线中有一条平行于 y 轴的．这条直线便是消去 y 后的方程．

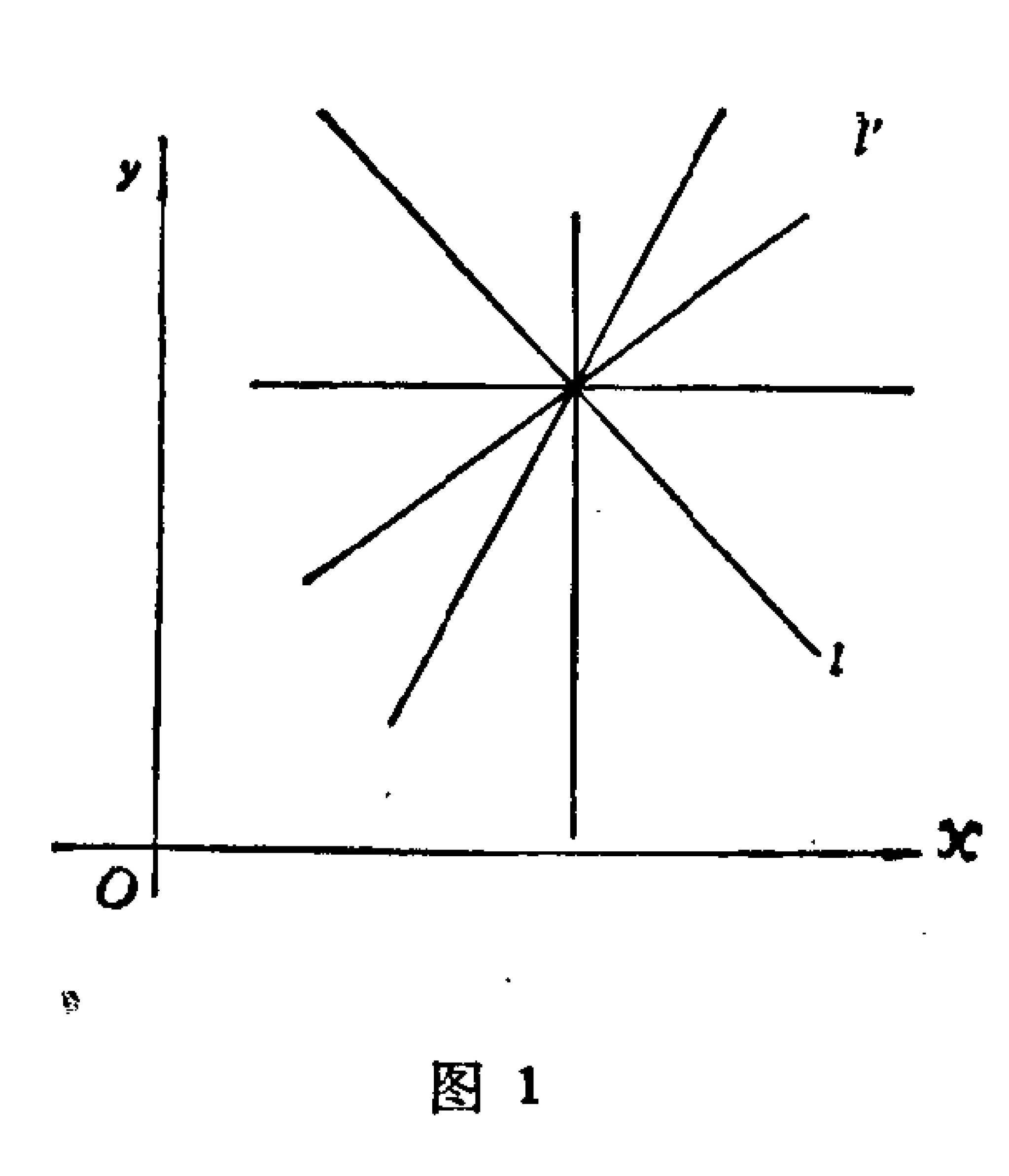


图 1

再看三个变数的情况．

l: $\quad ax+by+cz=d,$

l': $\quad a'x+b'y+c'z=d',$

l'': $\quad a''x+b''y+c''z=d''.$

这表示三个平面，平面族

$$\lambda l+\mu l'=0$$

代表通过 l 与 l' 的交线的所有的平面．由 l 与 l' 中消去 x 而得出的方程可以看成为：它代表通过交线而平行于 x 轴的平面，也可以看成为：这条交线在 (y,z) 平面上的投影，就是 y, z 平面上的一条直线，再从 l, l'' 中消去 x，又得 y, z 平面上的一条直线．

因此，消去 x，可以看作把三维空间的三个平面求交点的问题变为在 y, z 平面上求两条直线的交点的问题．这两条直线，正是两条空间直线(平面的交线)的投影．

一般讲来：一个

$$a_{11}x_1+\cdots+a_{1m}x_n=b_n$$

可以看成为 n 维空间的超平面．消去法便是把 n 从空间 m 个超平面求交点的问题化为 $n-1$ 维空间 $n-1$ 个超平面求交点的问题．

§4．消去法的力学解释

在一条两端固定的弦线上取 n 点 $P_1,\cdots,P_n$，在这 n 点各加一重物，也就是在这些点各有一向下的力 $F_1,\cdots,F_n$，我们来研究这些点的垂度 $y_1,\cdots,y_n$．

我们假定弦线上的力适合于“线性叠加原则”．

1°．两组力叠加，其对应的垂度也相加．

2°．所有力都乘以同一实数，则所有的垂度也乘上这一个相同的数．

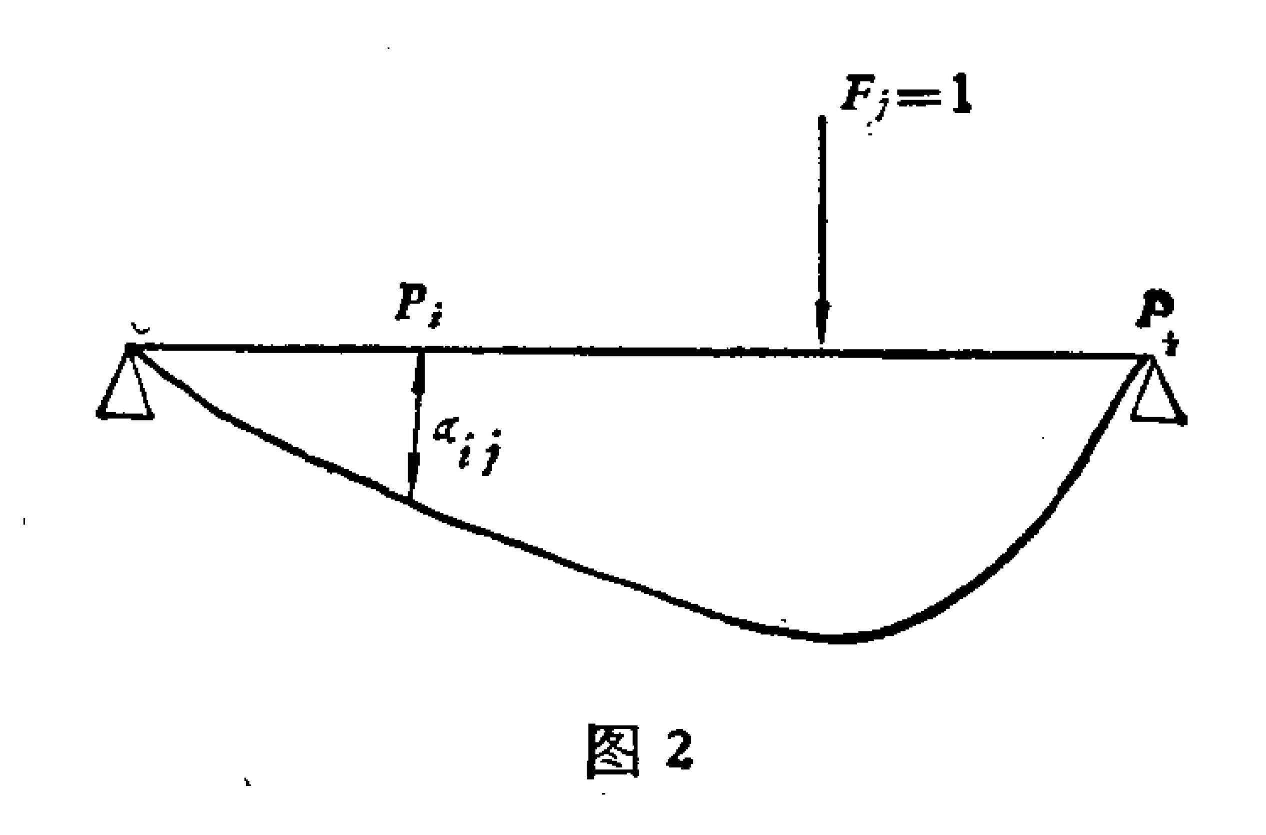


图 2

以 a_{ij} 表示当在 P_j 点上作用一个单位力时点 P_i 的垂度．这样，力 $F_1,\cdots,F_n$ 的联合作用后的垂度 $y_1,\cdots,y_n$ 等于

$$\sum_{j=1}^{n}a_{ij}F_j=y_i\quad(i=1,2,\cdots,n).\tag{1}$$

解线性方程组的问题，也就是给了垂度 $y_1,\cdots,y_n$ 要求出力 $F_1,\cdots,F_n$ 的问题了．

在 P_1 点加一个反作用力 R，这样单位力作用于 P_j 时，P_i 点的垂度等于

$$b_{ij}=a_{ij}+Ra_{i1}.$$

考虑把弦线固定于 P_1 的情况，也就是

$$b_{1j}=0,\qquad a_{1j}+Ra_{11}=0,$$

$$R=-a_{1j}/a_{11}.$$

也就是在 P_j 点加一个单位力，如果要在 P_1 加一个力使 P_1 固定，这个力是 $-a_{1j}/a_{11}$，这时

$$b_{ij}=a_{ij}-a_{1j}a_{i1}/a_{11}.$$

从 (1) 式消去 F_1，得

$$\sum_{j=2}^{n}(a_{ij}-a_{1j}a_{i1}/a_{11})F_j=y_i-a_{i1}y_1/a_{11}. \tag{2}$$

这就是加支点后的平衡方程，在 P_1 加了支点，在 P_j 作用一个单位力，b_{ij} 就是 P_i 的垂度。逐步消去，就是逐步加支点的过程。

§5. 经 济 平 衡

假定有 n 种生产品 $P_1,\cdots,P_n$ 生产一个单位 P_j 需要 a_{ij} 单位 P_i，如果各产品的数量是 $x_1,\cdots,x_n$，为了生产这些产品，P_i 类产品的总消耗是

$$\sum_{j=1}^{n}a_{ij}x_j,\quad 1\leqslant i\leqslant n.$$

能够供给市场的数量是

$$x_i-\sum_{j=1}^{n}a_{ij}x_j=b_i.$$

因此，知道了市场需要 $b_1,\cdots,b_n$，反回来考虑给各工业的生产指标 $x_1,\cdots,x_n$ 也是一个解线性方程的问题.

这类方程当然可以用消去法解，但更好是用叠代法解，关于叠代法将来再谈.

§6. 线性迴归分析

某一变量 ξ 决定于 n 个因素

$$\eta_1,\cdots,\eta_n.$$

我们已经做了 N 次实验得出的实验数据是

$$\xi^{(i)},\ \eta_1^{(i)},\ \eta_2^{(i)},\cdots,\ \eta_n^{(i)},\quad i=1,\cdots,N.$$

我们考虑线性关系

$$\xi=\sum_{j=1}^{n}a_j\eta_j$$

问题是怎样的线性关系，差方和最小，也就是，如果使

$$\zeta^{(i)}=\sum_{j=1}^{n}a_j\eta_j^{(i)},$$

求怎样的 a_j 使

$$\sum_{i=1}^{N}(\xi^{(i)}-\zeta^{(i)})^2$$

最小即求

$$F(a_1\cdots, a_n) = \sum_{i=1}^{N}\left(\xi^{(i)} - \sum_{j=1}^{n} a_j\eta_j^{(i)}\right)^2 \tag{1}$$

的极小值.

命

$$a_{jk} = \sum_{i=1}^{N}\eta_j^{(i)}\eta_k^{(i)}, \qquad b_k = \sum_{i=1}^{N}\xi^{(i)}\eta_k^{(i)}.$$

我们现在证明

$$\sum_{j=1}^{n} a_{jk}x_j = b_k, \qquad k = 1, 2, \cdots, n \tag{2}$$

的解答 $x_j = a_j$ 使 (1) 取最小值.

我们现在来证明这一点：如果 $a_1,\cdots,a_n$ 并不适合于(2). 例如：有一个 k 使

$$\sum_{j=1}^{n} a_{jk}a_j - b_k = -\alpha_k \neq 0.$$

我们考虑

$$\begin{aligned}
&F(a_1,\cdots,a_{k-1},a_k+\varepsilon,a_{k+1},\cdots,a_n)\\
&= \sum_{i=1}^{N}\left(\xi^{(i)} - \sum_{i=1}^{n} a_j\eta_j^{(j)} + \varepsilon\eta_k^{(i)}\right)^2\\
&= \sum_{i=1}^{N}\left(\xi^{(i)} - \sum_{j=1}^{n} a_j\eta_j^{(i)}\right)^2 + 2\varepsilon\sum_{i=1}^{N}\left(\xi^{(i)} - \sum_{j=1}^{n} a_i\eta_j^{(i)}\right)\eta_k^{(i)} + \varepsilon^2\sum_{i=1}^{n}(\eta_k^{(i)})^2\\
&= \sum_{i=1}^{N}\left(\xi^{(i)} - \sum_{j=1}^{n} a_j\eta_i^{(i)}\right)^2 + 2\varepsilon\left(b_k - \sum_{j=1}^{n} a_{jk}a_j\right) + \varepsilon^2 a_{kk}\\
&= F(a_1,\cdots,a_k) + 2\varepsilon\alpha_k + \varepsilon^2 a_{kk}.
\end{aligned}$$

凑方得

$$\begin{aligned}
&F(a_1,\cdots, a_{k-1}, a_k+\varepsilon, a_k,\cdots, a_n)\\
&\qquad = F(a_1, a_2,\cdots,a_n) + a_{kk}\left(\varepsilon + \frac{\alpha_k}{a_{kk}}\right)^2 - \frac{\alpha_k^2}{a_{kk}}.
\end{aligned} \tag{3}$$

如果 $\alpha_k \neq 0$，则 $F(a_1,\cdots,a_n)$ 不是最小值，因为在 (3) 式中取 $\varepsilon = -\alpha_k|a_{kk}$，则

$$F(a_1,\cdots,a_{k-1}, a_k+\varepsilon,a_k,\cdots,a_n)$$

的数值小于 $F(a_1,\cdots,a)$ 的数值了.

因此：求迴归平面的问题一变而为解线性方程的问题了.

致于要证明，适合于 (2) 的解一定使 F 取最小值，这一点的证明不难，如果 (2) 仅有一个解，当然毫无问题，因为由 (3) 可知不适合 (2) 的都不可能使 F 极小.（读者自证：(2)一定有解，并处理 (2) 有不止一个解的情况.）

方程组 (2) 当然可以用消去法来解，但是这是一个有对称系数的联立方程式，即

$$a_{jk} = a_{kj},$$

关于这样的方程组我们另有较好的计算方法.

以上的证明的优点之一，也许有人会指出，它避开了微积分，直接用初等的"凑方"法来处理了，实际上，更好的优点在于这个方法介绍了计算数学上的一个重要方法——松弛法.

特别在计算迴归分析时，松弛法更有价值，方法是：

1) 先任意地取一组 $a_1, \cdots, a_n$.

2) 任意地算一个

$$\alpha_k = \sum_{j=1}^{n} a_{jk} a_j - b_k,$$

如果 $\alpha_k \neq 0$，把

$$a_1, \cdots, a_{k-1}, a_k + \varepsilon, \cdots, a_n$$

作为原出发点，如果 $\alpha_k = 0$，则要换一个 k.

3) 一般的办法是 $k = 1, 2, \cdots, n, 1, 2, \cdots, n, \cdots$ 周而复始地进行计算，这样便可以得出所求的解答了.

这方法之所以命名为松弛法的原因固在于此，另一点是如果算错了，不必从头算，依错算下去，依然得出正确的结果来(即从错了的 $(a_1, \cdots, a_n)$ 再开始算下去，依然能得出结果来的).

当然，并不是说常常错，而是说偶然算错了关系不大而已.

虽然"松弛"，但偶而略为紧张些可以帮我们更有效地解决问题，例如：比较一下

$$\alpha_k / a_{kk}, \qquad k = 1, 2, \cdots, n$$

谁大，取使这值最大的整数 k 出发最有利，因为由(3)可知在 F 上减得多了，这方法一定可以逐步逼近原解答的.

§7. 行 列 式

建议从 §1 的关系引进行列式，也就是用数学归纳法来定义行列式，即行列式

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}.$$

此处 A_{ij} 是由原行列中划掉第 i 行，第 j 列所得出的 $n-1$ 行的行列式的数值，再乘以 $(-1)^{i+j}$.

A_{ij} 称为 a_{ij} 的余因子.

行列式的重要性质：

(1) 一行(列)同以 k 乘之，则行列式的数值是原来的 k 倍.

(2) 把一行(列)的 h 倍加到另一行(列)上，行列式的数值不变.

(3) 两行(列)互换，行列式变号，由此可知两行相等，行列式之值为 0.

解方程式的 Cramer 法则：

$$\sum_{j=1}^{n} a_{ij} x_j = b_i$$

的解答是

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}}{|a_{ij}|},$$

x_i 有相似的表达式，但这个是把 $|a_{ij}|$ 中的第 1 列换为"b"，而那个是把第 i 列换为"b"。

齐次方程的基本定理：

$$\sum_{j=1}^{n} a_{ij}x_j = 0, \quad i = 1, \cdots, n.$$

有非显见解的必要且充分条件是

$$|a_{ij}| = 0.$$

简单推论 1. 如果 $n > m$，方程组

$$\sum_{j=1}^{n} a_{ij}x_j = 0, \quad i = 1, \cdots, m$$

一定有非显见解，因为我们可以虚加上 $n-m$ 个方程

$$\sum_{j=1}^{n} a_{ij}x_j = 0, \quad i = m+1, \cdots, n.$$

其中 $a_{ij} = 0$，$m+1 \leqslant i \leqslant n$，$1 \leqslant j \leqslant n$.

简单推论 2. 如果 $n \leqslant m$，则方程组

$$\sum_{i=1}^{n} a_{ij}x_j = 0, \quad i = 1, \cdots, m \tag{1}$$

有非显见解，一定是其中任意 n 个方程的行列式都等于 0。

比基本定理较一般些的结果是，如果任意 n 个方程式的行列式都等于 0，则 (1) 有一个非显见解.

证明 1) 由假定可知

$$\begin{vmatrix} a_{11}, & \cdots, & a_{1n} \\ a_{22}, & \cdots, & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{n-11}, & \cdots, & a_{n-1n} \\ a_{i1}, & \cdots, & a_{in} \end{vmatrix} = 0.$$

把这式子展开得

$$a_{i1}A_1 + \cdots + a_{in}A_n = 0, \qquad A_n = \begin{vmatrix} a_{11}, & \cdots, & a_{1n-1} \\ \cdots & \cdots & \cdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{vmatrix}.$$

如果 $A_1, \cdots, A_n$ 不全等于 0，则方程组 (1) 显然有解 $x_i = A_i$.

2) 如果经过重排方程的次序，或重编 x_i 的号码，使

$$\begin{vmatrix} a_{11} & \cdots & a_{1n-1} \\ \cdots & \cdots & \cdots \\ a_{n-11} & \cdots & a_{n-1,n-1} \end{vmatrix} \neq 0,$$

则由 1) 可知本定理正确.

3) 现在考虑 2) 没有包括进去的情况. 考虑 $x_n = 0$ 时的情况，现在有

$$\begin{vmatrix} a_{11} & \cdots, & a_{1n-1} \\ \cdots & \cdots & \cdots \\ a_{n-21}, & \cdots, & a_{n-2n-1} \\ a_{i1}, & \cdots, & a_{in-1} \end{vmatrix} = 0.$$

展开得

$$a_{i1}B_1+\cdots+a_{in-1}B_{n-1}=0.$$

即取 $x_1=B_1,\cdots,x_{n-1}=B_{n-1},x_n=0$ 就是解.

这样可以运用归纳法来证明本定理.

现在来研究 §1 中可提出的齐次方程组与非齐次方程组的关系.

首先：如果

$$\sum_{j=1}^{n}a_{ij}x_j=0,\quad i=1,\cdots,n$$

有非显见解，则

$$\sum_{i=1}^{n}a_{ij}x_i=0,\quad j=1,\cdots,n \tag{2}$$

也有非显见解.

这是显然的,因为行换为列,行列式的数值不变. 命

$$\sum_{i=1}^{n}\xi_i a_{ij}=0,$$

并可假定 $\xi_1\neq 0$ 如此则由

$$\sum_{j=1}^{n}a_{ij}x_j=b_i \tag{3}$$

可知

$$\sum_{i=1}^{n}b_i\xi_i=\sum_{i=1}^{n}\xi_i\left(\sum_{j=1}^{n}a_{ij}x_j\right)=\sum_{j=1}^{n}\left(\sum_{i=1}^{n}a_{ij}\xi_i\right)x_j=0.$$

显然 $b_1=1,b_2=\cdots=b_n=0$ 时,方程组(3)无解.

如果 (2) 仅有显见解,则

$$|a_{ij}|\neq 0.$$

由 Cramer 公式可知方程 (2) 有解.

注意 Cramer 公式虽然漂亮,但是真正解方程式时不常用它,因为其中运算的次数太多了.

§8. Vandermonde 行列式

定理 1

$$\Delta=\Delta(x_1,x_2,\cdots x_n)=\begin{vmatrix}1, 1, & \cdots & 1\\ x_1, x_2, & \cdots, & x_n\\ \cdots\cdots\cdots \\ x_1^{n-1}, x_2^{n-1}, & \cdots, & x_n^{n-1}\end{vmatrix}=\prod_{n\geq i>j\geq 1}(x_i-x_j).$$

证明 1) 用归纳法, $n=2$ 时,显然正确.

2) 在第 2, 3,…, n 行中各减前一行的 x_1 倍如此得

$$\Delta=\begin{vmatrix}x_2-x_1,\cdots,x_n-x_1\\ x_2(x_2-x_1),\cdots,x_n(x_n-x_1)\\ \cdots\cdots\cdots\cdots \\ x_2^{n-2}(x_2-x_1),\ \cdots,\ x_n^{n-2}(x_n-x_1)\end{vmatrix}$$

$$= (x_2 - x_1)\cdots(x_n - x_1)\begin{vmatrix} 1, & \cdots\cdots, & 1 \\ x_2, & \cdots\cdots, & x_n \\ \cdots & \cdots & \cdots \\ x_2^{n-2}, & \cdots\cdots, & x_2^{n-2} \end{vmatrix},$$

用归纳法即得所求.

附证 当 $x_i = x_j$ 时 $\Delta = 0$, 因此 $x_i - x_j$ 可除尽 Δ, 因此

$$\prod_{n \geq i > j \geq 1} (x_i - x_j)$$

可除尽 Δ, 再比较 $x_n^{n-1} x_{n-1}^{n-2} \cdots x_2 \cdot 1$ 的系数可得本定理.

这定理有以下的显然推广.

定理 2 命 $P_i(x)$ 是第 i 次多项式, 其 x^i 的系数是 a_i. 如此则

$$\begin{vmatrix} P_0(x_1), & \cdots, & P_0(x_n) \\ P_1(x_1), & \cdots, & P_1(x_n) \\ \cdots & \cdots & \cdots \\ P_{n-1}(x_1), & \cdots, & P_{n-1}(x_n) \end{vmatrix} = a_0 \cdots a_n \prod_{n \geq i > j \geq 1} (x_i - x_j). \qquad (1)$$

这个结果的证明是容易的, 首先, 第一行全是 a_0, 以 a_0 除这一行, 得一同为 1 的行的行列式, 命

$$P_2(x) = a_1 x + a_1'$$

在第二行中减去第一行的 a_1' 倍, 再除以 a_1, 则第二行变为

$$x_1, \cdots\cdots, x_n.$$

命

$$P_2(x) = a_2 x^2 + a_2' x + a_2''$$

在第三行中减去第一行的 a_2'' 倍, 减去第二行的 a_2' 倍, 再除以 a_2, 第三行变为

$$x_1^2, \cdots\cdots, x_n^2.$$

依此绕行, 即得所求的公式了.

定理 3 我们有

$$\begin{vmatrix} 1, & \cdots, & 1 \\ \cos\theta_1, & \cdots, & \cos\theta_n \\ \cdots & \cdots & \cdots \\ \cos(n-1)\theta_1, & \cdots, & \cos(n-1)\theta_n \end{vmatrix} = 2^{\frac{1}{2}(n-2)(n-1)} \prod_{n \geq i > j \geq 1} (\cos\theta_i - \cos\theta_j)$$

及

$$\begin{vmatrix} \sin\theta_1, & \cdots, & \sin\theta_n \\ \sin 2\theta_1, & \cdots, & \sin 2\theta n \\ \cdots & \cdots & \cdots \\ \sin n\theta_1, & \cdots, & \sin n\theta_n \end{vmatrix} = 2^{\frac{1}{2}n(n-1)} \sin\theta_1 \cdots \sin\theta_n \prod_{n \geq i > j \geq 1} (\cos\theta_i - \cos\theta_j).$$

证明 由于

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta,$$

因此

$$2\cos n\theta = (\cos\theta + i\sin\theta)^n + (\cos\theta - i\sin\theta)^n$$

$$= \sum_{l=0}^{n} \binom{n}{l} [(i\sin\theta)^l + (-i\sin\theta)^l]\cos^{n-l}\theta$$

$$= 2\sum_{k=0}^{[n/2]} \binom{n}{2k} (-1)^k \sin^{2k}\theta\cos^{n-2k}\theta.$$

即

$$\cos n\theta = \sum_{k=0}^{[n/2]} (-1)^k \binom{n}{2k} (1-\cos^2\theta)^k \cos^{n-2k}\theta = P_n(\cos\theta).$$

这儿 P_n 是 $\cos\theta$ 的多项式其中 $\cos^n\theta$ 的系数等于

$$\sum_{k=0}^{[n/2]} (-1)^k(-1)^k \binom{n}{2k} = \frac{1}{2}\sum_{l=0}^{n} (1+(-1)^l)\binom{n}{l}$$
$$= 2^{n-1}.$$

以此代入原行列式，得

$$\begin{vmatrix} 1, & \cdots, & 1 \\ \cos\theta_1, & \cdots, & \cos\theta_n \\ \cdots & \cdots & \cdots \\ \cos(n-1)\theta_1, & \cdots, & \cos(n-1)\theta_n \end{vmatrix} = \begin{vmatrix} P_0(\cos\theta_1), & \cdots, & P_0(\cos\theta_n) \\ P_1(\cos\theta_1), & \cdots, & P_1(\cos\theta_n) \\ \cdots & \cdots & \cdots \\ P_{n-1}(\cos\theta_1), & \cdots, & P_{n-1}(\cos\theta_n) \end{vmatrix}$$

$$= 2^{0+1+2+\cdots+(n-2)} \prod_{n\geqslant i>j\geqslant 1} (\cos\theta_i - \cos\theta_j)$$

又

$$2i\sin n\theta = (\cos\theta + i\sin\theta)^n - (\cos\theta - i\sin\theta)^n$$

$$= \sum_{l=0}^{n} \binom{n}{l} [(i\sin\theta)^l - (i\sin\theta)^l]\cos^{n-l}\theta$$

$$= 2i\sum_{0\leqslant 2k+1\leqslant n} \binom{n}{2k+1} (-1)^k \sin^{2k+1}\theta\cos^{n-2k-1}\theta.$$

即

$$\sin n\theta = \sin\theta \sum_{0\leqslant 2k+1\leqslant n} \binom{n}{2k+1} (-1)^k (1-\cos^2\theta)^k \cos^{n-2k-1}\theta$$

$$= \sin\theta Q_{n-1}(\cos\theta).$$

这儿 $Q_{n-1}(\cos\theta)$ 是 $\cos\theta$ 的 $n-1$ 次多项式,其 $\cos^{n-1}\theta$ 的系数为

$$\sum_{0\leqslant 2k+1\leqslant n} \binom{n}{2k+1} = \sum_{l=0}^{n} \binom{n}{l} - \sum_{0\leqslant 2k\leqslant n} \binom{n}{2k} = 2^n - 2^{n-1} = 2^{n-1}.$$

因此

$$\begin{vmatrix} \sin\theta_1, & \cdots, & \sin\theta_n \\ \sin 2\theta_1, & \cdots & \sin 2\theta_n \\ \cdots & \cdots & \cdots \\ \sin n\theta_1, & \cdots, & \sin n\theta_n \end{vmatrix} = \sin\theta_1\cdots\sin\theta_n \begin{vmatrix} Q_0(\cos\theta), & \cdots, & Q_0(\cos\theta_n) \\ \cdots & \cdots & \cdots \\ Q_{n-1}(\cos\theta_1), & \cdots, & Q_{n-1}(\cos\theta_n) \end{vmatrix}$$

$$= 2^{\frac{1}{2}(n-1)n}\sin\theta_1\cdots\sin\theta_n \prod_{n\geqslant i>j\geqslant 1} (\cos\theta_i - \cos\theta_j).$$

定理 4

$$\begin{vmatrix} \sin\frac{1}{2}\theta_1, & \cdots, & \sin\frac{1}{2}\theta_n \\ \sin\frac{3}{2}\theta_1 & \cdots, & \sin\frac{3}{2}\theta_n \\ \cdots & \cdots & \cdots \\ \sin\left(n-\frac{1}{2}\right)\theta_n, & \cdots, & \sin\left(n-\frac{1}{2}\right)\theta_n \end{vmatrix} = 2^{\frac{1}{2}n(n-1)}\prod_{i=1}^{n}\sin\frac{1}{2}\theta_i\prod_{n\geq i>j\geq 1}(\cos\theta_i-\cos\theta_j).$$

证明　第 1,2,…, n 列各乘以 $\cos\frac{1}{2}\theta_1\cdots,\cos\frac{1}{2}\theta_n$ 由于:

$$\sin\left(v-\frac{1}{2}\right)\theta_i\cos\frac{1}{2}\theta_i=\frac{1}{2}(\sin v\theta_i+\sin(v-1)\theta_i),$$

所以原行列式等于

$$\frac{1}{2^n\prod_{i=1}^{n}\cos\frac{1}{2}\theta_i}\begin{vmatrix} \sin\theta_1, & \sin\theta_2, & \cdots, & \sin\theta_n \\ \sin 2\theta_1-\sin\theta_1, & \sin 2\theta_2-\sin\theta_2, & \cdots, & \sin 2\theta_n-\sin\theta_n \\ \cdots & \cdots & \cdots & \cdots \\ \sin n\theta_1-\sin(n-1)\theta_2, & & \cdots & \sin n\theta_n-\sin(n-1)\theta_n \end{vmatrix}$$

$$=\frac{1}{2^n\prod_{i=1}^{n}\cos\frac{1}{2}\theta_i}\begin{vmatrix} \sin\theta_1, \sin\theta_2, & \cdots, & \sin\theta_n \\ \sin 2\theta_1, \sin 2\theta_2, & \cdots, & \sin 2\theta_n \\ \cdots & \cdots & \cdots \\ \sin n\theta_1, \sin n\theta_n, & \cdots, & \sin n\theta_n \end{vmatrix}$$

$$=2^{\frac{1}{2}n(n-1)}\sin\frac{1}{2}\theta_1\cdots\sin\frac{1}{2}\theta_n\prod_{n\geq i>j\geq 1}(\cos\theta_i-\cos\theta_j).$$

定理 5

$$\begin{vmatrix} \cos\frac{1}{2}\theta_1, & \cdots, & \cos\frac{1}{2}\theta_n \\ \cos\frac{3}{2}\theta_1, & \cdots, & \cos\frac{3}{2}\theta_n \\ \cdots & \cdots & \cdots \\ \cos\left(n-\frac{1}{2}\right)\theta_1, & \cdots, & \cos\left(n-\frac{1}{2}\right)\theta_n \end{vmatrix} = 2^{\frac{1}{2}n(n-1)}\prod_{i=1}^{n}\cos\frac{1}{2}\theta_i\prod_{n\geq i>j\geq 1}(\cos\theta_i-\cos\theta_j).$$

证明　第 1, 2,… n 列是各乘以

$$\sin\frac{1}{2}\theta_1,\ \cdots\sin\frac{1}{2}\theta_n.$$

由于

$$\sin\frac{1}{2}\theta_i\cos\left(v+\frac{1}{2}\right)\theta_i=\frac{1}{2}(\sin(v+1)\theta_i-\sin v\theta_i),$$

所以原行列式等于

$$\frac{1}{2^n\prod_{i=1}^{n}\sin\frac{1}{2}\theta_i}\begin{vmatrix} \sin\theta_1, & \cdots, & \sin\theta_n \\ \sin 2\theta_1-\sin\theta_1, & \cdots, & \sin 2\theta_n-\sin\theta_n \\ \sin 3\theta_2-\sin 2\theta_1, & \cdots, & \sin 2\theta_n-\sin\theta_n \\ \cdots & \cdots & \cdots \\ \sin n\theta_1-\sin(n-1)\theta_1, & \cdots & \sin n\theta_n-\sin(n-1)\theta_n \end{vmatrix}$$

$$= \frac{1}{2^n \prod_{i=1}^{n} \sin \frac{1}{2}\theta_i} \begin{vmatrix} \sin\theta_1, & \cdots, & \sin\theta_n \\ \sin 2\theta_1, & \cdots, & \sin 2\theta_n \\ \cdots & \cdots & \cdots \\ \sin n\theta_n, & \cdots, & \sin n\theta_n \end{vmatrix}.$$

$$= 2^{\frac{1}{2}(n-1)n} \cos \frac{1}{2}\theta_1 \cdots \cos \frac{1}{2}\theta_n \prod_{n \geqslant i > j \geqslant 1} (\cos\theta_i - \cos\theta_j).$$

习题 1　用某一次序来标列一个多边体的 n 个顶点，现在做以下的行列式，如果第 i 个顶点和第 j 个顶点之中有一边相联，则取 $a_{ij} = a_{ji} = 1$。 不然则取 $a_{ij} = 0$，特别有 $a_{ii} = 0$。

证明　这个行列式的数值与顶点排列的次序无关，并且算出四面体，六面体，八面体的行列式：例如四面体

$$\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} = -3.$$

提示　把第 i 点换为第 j 点，则第 i 行与第 j 行，第 i 列与第 j 列同时进行交换.

答数　六面体：9，八面体：0.

习题 2　算出

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ b_1, & a_1, & a_1 & \cdots & a_1 \\ b_1, & b_2, & a_2 & \cdots & a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_1, & b_2, & b_3 & \cdots & a_n \end{vmatrix}.$$

习题 3　试证

$$\begin{vmatrix} \rho & \frac{l}{m} + \frac{l}{m} & \frac{n}{l} + \frac{l}{n} \\ \frac{l}{m} + \frac{m}{l} & \rho & \frac{m}{n} + \frac{n}{m} \\ \frac{n}{l} + \frac{l}{n} & \frac{m}{n} + \frac{n}{m} & \rho \end{vmatrix}, \quad l \neq 0, \ m \neq 0, \ n \neq 0$$

$$= (\rho - 2)(\rho + 1 + \sqrt{P})(\rho + 1 - \sqrt{P}).$$

此处

$$P = \left(\frac{1}{l^2} + \frac{1}{m^2} + \frac{1}{n^2}\right)(l^2 + m^2 + n^2).$$

§9. 对 称 函 数

把乘积

$$f(x) = (x - x_1)\cdots(x - x_n) \tag{1}$$

展开成为

$$f(x) = x^n - \sigma_1 x^{n-1} + \sigma_2 x^{n-2} - \sigma_3 x^{n-3} + \cdots. \qquad (2)$$

这儿

$$\sigma_1 = x_1 + x_2 + \cdots + x_n,$$

$$\sigma_2 = x_1x_2 + x_1x_3 + \cdots x_1x_n + x_2x_3 + \cdots + x_2x_n + \cdots\cdots\cdots + x_{n-1}x_n.$$

其中 σ_i 是所有的从 $x_1, \cdots, x_n$ 中取 i 个乘积的总和，这个表达式中有 $\binom{n}{i}$ 项，每项是 i 个数的乘积.

这些 $\sigma_1, \cdots, \sigma_n$ 称为 $x_1, \cdots, x_n$ 的初等对称函数.

定义　如果把 $x_1, \cdots, x_n$ 重排成为 $x_{i_1}, \cdots, x_{i_n}$，函数

$$F(x_1, \cdots, x_n) = F(x_{i_1}, \cdots, x_{i_n}),$$

如果所有的重排这关系式都成立，则函数 F 称为 $x_1, \cdots, x_n$ 的对称函数.

初等对称函数是对称函数，还另有一主要的对称函数是对称幂和，即

$$s_i = x_1^i + \cdots + x_n^i.$$

定义 $\sigma_0 = 1, s_0 = n$. 取出一项 $x_{i_1}^{l_1} x_{i_2}^{l_2} \cdots x_{i_m}^{l_m}$，用

$$\sum x_{i_1}^{l_1} x_{i_2}^{l_2} \cdots x_{i_m}^{l_m}$$

表示由该项经由可能的排列所得出的总和，例如：

$$\sigma_2 = \sum x_i x_j,$$

注意这不同于

$$\sum_{j=1}^{n} \sum_{i=1}^{n} x_i x_j.$$

因为一个没有 x_i^2 项，而一个有.

在 (2) 中代以 x_i，得

$$x_i^n - \sigma_1 x_i^{n-1} + \sigma_2 x_i^{n-2} - \cdots + (-1)^n \sigma_n = 0, \quad i = 1, 2, \cdots, n. \qquad (3)$$

把这看成为 n 个线性方程式，把 $-\sigma_1, \sigma_2, \cdots, (-1)^n\sigma_n$ 看成为未知数. 用 Cramer 公式解得

$$(-1)^{n-l}\sigma_{n-l} = \begin{vmatrix} x_1^{n-1}, & \cdots, & x_1^{l+1}, -x_1^n, x_1^{l-1}, & \cdots, & x_1, 1 \\ x_2^{n-1}, & \cdots, & x_n^{l+1}, -x_2^n, x_2^{l-1}, & \cdots, & x_2, 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_n^{n-1}, & \cdots, & x_n^{l+1}, -x_n^n, x_n^{l-1}, & \cdots, & x_n, 1 \end{vmatrix} \Bigg/ \begin{vmatrix} x_1^{n-1}, x_1^{n-2}, & \cdots, & x_1, 1 \\ x_2^{n-1}, x_2^{n-2}, & \cdots, & x_2, 1 \\ \cdots & \cdots & \cdots \\ x_n^{n-1}, x_n^{n-2}, & \cdots, & x_n, 1 \end{vmatrix}. \qquad (4)$$

因此得出

定理 1　命

$$\Delta(l_1, \cdots, l_n) = \begin{vmatrix} x_1^{l_1}, x_2^{l_1}, & \cdots, & x_n^{l_1} \\ \cdots & \cdots & \cdots \\ x_1^{l_n}, x_2^{l_n}, & \cdots, & x_n^{l_n} \end{vmatrix},$$

则

$$\sigma_{n-l} = \frac{\Delta(n, n-1, \cdots, l+1, l-1, \cdots, 0)}{\Delta(n-1, n-2, \cdots, 1, 0)}.$$

习题 1　证明 $\Delta(n-1, n-2, \cdots, 1, 0)$ 除得尽所有的 $\Delta(l_1, \cdots, l_n)$，并且其商是对称函数.

习题 2　试算出

$$\Delta(n+1, n-2, \cdots, 1, 0),$$
$$\Delta(n+1, n, n-2, n-3\cdots, 1, 0).$$

提示：先证明形式上有

$$\Delta(n+1, n-2, \cdots, 1, 0) = \Delta(n-1, n-2, \cdots, 1, 0) \times (A\sum x_i^2 + B\sum x_i x_j).$$

再定出 A 与 B. 我们现在研究

$$\sigma_1, \sigma_2, \cdots, \sigma_n$$

与

$$s_1, s_2, \cdots, s_n$$

的关系由式 (1) 的对数微商可知

$$f(x) = \frac{f(x)}{x-x_1} + \frac{f(x)}{x-x_2} + \cdots + \frac{f(x)}{x-x_n}.$$

由于

$$\frac{x^l - x_1^l}{x - x_1} = x^{l-1} + x^{l-2}x_1 + \cdots + x_1^{l-1},$$

可知

$$\frac{f(x)}{x-x_i} = \frac{f(x)-f(x_i)}{x-x_i} = \sum_{l=0}^{n}(-1)^l\sigma_l\frac{x^{n-l}-x_i^{n-l}}{x-x_i}$$
$$= \sum_{l=0}^{n-1}(-1)^l\sigma_l\sum_{t=1}^{n-l-1}x^t x_i^{n-l-t-1}$$
$$= \sum_{t=0}^{n-1}\left(\sum_{l=0}^{n-t-1}(-1)^l\sigma_l x_i^{n-l-t-1}\right)x^t.$$

对 i 求和因此得出

$$f'(x) = \sum_{t=0}^{n-1}\left(\sum_{l=0}^{n-t-1}(-1)^l\sigma_l s_{n-l-t-1}\right)x^t.$$

另一方面，微分 (2) 式得

$$f'(x) = \sum_{t=0}^{n}(-1)^{n-t}\sigma_{n-t}(x^t)' = \sum_{t=1}^{n}(-1)^{n-t}t\sigma_{n-t}x^{t-1}.$$

比较系数得

$$\sum_{l=0}^{n-t-1}(-1)^l\sigma_l s_{n-l-t-1} = (-1)^{n-t-1}(t+1)\sigma_{n-t-1}.$$

$$t = 0, 1, 2, \cdots, n-1.$$

把 t 换为 $n-\tau$，则得

$$\sum_{l=0}^{\tau-1}(-1)^l\sigma_l s_{\tau-l-1} = (-1)^{\tau-1}(n-\tau+1)\sigma_{\tau-1},$$

$$\tau = 1, 2, \cdots, n-1, n,$$

$$s_0 = n\sigma_0,$$
$$\sigma_0 s_1 - \sigma_1 s_0 = -(n-1)\sigma_1,$$
$$\sigma_0 s_2 - \sigma_1 s_1 + \sigma_2 s_0 = (n-2)\sigma_2,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\sigma_0 s_m - \sigma_1 s_{m-1} + \cdots + (-1)^m \sigma_m s_0 = (-1)^m (n-m)\sigma_m.$$

以 $\sigma_0 = 1$, $s_0 = n$ 代入,并移项可得

$$s_1 = \quad = \sigma_1,$$
$$\sigma_1 s_1 - s_2 = 2\sigma_2,$$
$$\cdots\cdots\cdots$$
$$\sigma_{m-1} s_1 - \sigma_{m-2} s_2 + \cdots + (-1)^{m-1} s_m = m\sigma_m.$$

把 $s_1, \cdots s_m$ 看成为未知数,解这线性方程组得:

定理 2 (Newton)

$$s_1 = \sigma_2,\ s_2 = \begin{vmatrix} \sigma_1 & 1 \\ 2\sigma_2 & \sigma_1 \end{vmatrix}, \quad s_3 = \begin{vmatrix} \sigma_1 & 1 & 0 \\ 2\sigma_2 & \sigma_1 & 1 \\ 3\sigma_3 & \sigma_2 & \sigma_1 \end{vmatrix}, \cdots$$

$$s_m = \begin{vmatrix} \sigma_1 & 1 & 0 \cdots \\ 2\sigma_2 & \sigma_1 & 1 \cdots \\ 2\sigma_3 & \sigma_2 & \sigma_1 \cdots \\ \cdots & \cdots & \cdots \\ m\sigma_m & \sigma_{m-1} & \sigma_{m-2} \cdots \end{vmatrix}, \quad m = 1, 2, \cdots, n.$$

再把 $\sigma_1, \cdots \sigma_m$ 看成为未知数解得:

定理 3

$$\sigma_2 = \frac{1}{2}\begin{vmatrix} s_1 & 1 \\ s_2 & s_1 \end{vmatrix}, \quad \sigma_3 = \frac{1}{3!}\begin{vmatrix} s_1 & 1 & 0 \\ s_2 & s_1 & 2 \\ s_3 & s_2 & s_1 \end{vmatrix}, \quad \sigma_4 = \frac{1}{4!}\begin{vmatrix} s_1 & 1 & 0 & 0 \\ s_2 & s_1 & 2 & 0 \\ s_3 & s_2 & s_1 & 3 \\ s_4 & s_3 & s_2 & s_1 \end{vmatrix}, \cdots.$$

如果 $m > n$, 则由

$$x_i^n - \sigma_1 x_i^{n-1} + \cdots + (-1)^n \sigma_n = 0$$

乘以 x_i^{m-n}, 再对 $i = 1, 2, \cdots, n$ 相加, 如此得出

$$s_m - \sigma_1 s_{m-1} + \cdots + (-1)^n \sigma_n s_{m-n} = 0,$$

即 s_m 可由 $\sigma_1, \cdots, \sigma_n$, 及 $s_{m-1}, \cdots, s_{m-n}$ 看出来,这是递归公式.

§ 10. 对称函数的基本定理

定理 1 任何一个对称函数(多项式)一定可以表为初等对称函数的函数.

证明 1) 如果能够证明,它可以表为 $s_1, s_l, \cdots, s_m, \cdots$ 的函数, 则由上节的结果知道,它可以表为 $\sigma_1, \cdots, \sigma_m$ 的函数.

2) 由

$$s_m = \sum_{i=1}^{n} x_i^m, \quad s_p = \sum_{i=1}^{n} x_i^p.$$

可得

$$s_m s_p = \sum x_i^{m+p} + \sum_{i \neq j} x_i^m x_j^p.$$

当 $m \neq p$ 时，可知对称函数

$$\sum x_i^m x_j^p = s_m s_p - s_{m+p}.$$

当 $m = p$ 时，

$$s_m s_m = s_{2m} + 2\sum x_i^m x_j^m.$$

即对称函数

$$\sum (x_i x_j)^m = \frac{1}{2}(s_m^2 - s_{2m}).$$

3) 再考虑有三个不同因子的对称函数 $\sum x_i^m x_j^p x_k^q$.

如果 m, p, q, 各不相等，且 $m + q \neq p$, $p + q \neq m$ 时，

$$s_q \sum x_i^m x_j^p = \sum x_i^{m+q} x_j^p + \sum x_i^m x_j^{p+q} + \sum x_i^m x_j^p x_k^q.$$

由 (2) 可知

$$s_q(s_m s_p - s_{m+p}) = s_{m+q} s_p - s_{m+p+q} + s_m s_{p+q} - s_{m+p+q} + \sum x_i^m x_j^p x_k^q.$$

因此

$$\sum x_i^m x_j^p x_k^q = s_m s_p s_q - s_{m+p} s_q - s_{m+q} s_p - s_{m+p} s_m + 2 s_{m+p+q}.$$

如果 m, p, q 各不相等，但 $m + q = p$ 时，则得

$$\sum x_i^m x_j^p x_k^q = s_m s_p s_q - s_q s_{m+p} - \frac{1}{2} s_p^2 - s_{p+q} s_m + \frac{3}{2} s_{2p}.$$

如果 m, p, q 各不相等，但 $p + q = m$ 时，则得

$$\sum x_i^m x_j^p x_k^q = s_m s_p s_q - s_q s_{m+p} - \frac{1}{2} s_m^2 - s_{m+q} s_p + \frac{3}{2} s_{2m}.$$

如果 $m = p$ 则得

$$2\sum (x_i x_j)^m x_k^q = s_m^2 s_q - s_{2m} s_q - 2 s_{m+q} s_m + 2 s_{2m+q}.$$

又如果 $m = p = q$, 即

$$6\sum (x_i x_j x_k)^m = s_m^3 - 3 s_{2m} s_m + 2 s_{3m}.$$

即 $\sum x_i^m x_j^p x_k^q$ 可以成为初等对称函数. 由此方法续行，可以证明本定理.

§11. 两个代数方程有无公根

以前讲过了一个重要原则: 如果有一组非全为 0 的数 $x_1, \cdots, x_n$ 使得

$$\sum_{j=1}^{n} a_{ij} x_j = 0, \quad i = 1, 2 \cdots, n,$$

则

$$|a_{ij}| = 0.$$

这一原则十分重要，并且是一个经常用的工具，现在举几个例子来说明此法的应用.

例 1　假定

$$a_0x^4 + a_1x^3 + a_2x^2 + a_3x + a_4 = 0, \tag{1}$$

$$b_0x^3 + b_1x^2 + b_2x + b_3 = 0, \tag{2}$$

有一个公根 ξ. 则

$$\begin{aligned}
a_0\xi^6 + a_1\xi^5 + a_2\xi^4 + a_3\xi^3 + a_4\xi^2 &= 0,\\
a_0\xi^5 + a_1\xi^4 + a_2\xi^3 + a_3\xi^2 + a_4\xi &= 0,\\
a_0\xi^4 + a_1\xi^3 + a_2\xi^2 + a_3\xi + a_4 &= 0,\\
b_0\xi^6 + b_1\xi^5 + b_2\xi^4 + b_3\xi^3 &= 0,\\
b_0\xi^5 + b_1\xi^4 + b_2\xi^3 + b_3\xi^2 &= 0,\\
b_0\xi^4 + b_1\xi^3 + b_2\xi^2 + b_3\xi &= 0,\\
b_0\xi^3 + b_2\xi^2 + b_2\xi + b_3 &= 0,
\end{aligned}$$

可以看为七个线性方程，有一个解答

$$x_1 = \xi^6,\ x_2 = \xi^5,\ x_3 = \xi^4,\ x_4 = \xi^3,\ x_3 = \xi^2,\ x_2 = \xi,\ x_7 = 1.$$

这是不全等于 0 的解答，因此

$$\Delta = \begin{vmatrix}
a_0 & a_1 & a_2 & a_3 & a_4 & 0 & 0\\
0 & a_0 & a_1 & a_2 & a_3 & a_4 & 0\\
0 & 0 & a_0 & a_1 & a_2 & a_3 & a_4\\
b_0 & b_1 & b_2 & b_3 & 0 & 0 & 0\\
0 & b_0 & b_1 & b_2 & b_3 & 0 & 0\\
0 & 0 & b_0 & b_1 & b_2 & b_3 & 0\\
0 & 0 & 0 & b_0 & b_1 & b_2 & b_3
\end{vmatrix} = 0.$$

这便是方程 (1)，(2) 有公根的条件了.

一般讲来

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n = 0,$$

$$g(x) = b_0x^m + b_1x^{m-1} + \cdots + b_m = 0,$$

有公根的条件可以由

$$x^{m-1}f(x) = 0, \cdots, xf(x) = 0, f(x) = 0,$$

$$x^{n-1}g(x) = 0, \cdots, xg(x)0 = 0, g(x) = 0,$$

消去 $x^{m+n-1}, \cdots, x, 1$ 而得出的行列式

$$\Delta = 0.$$

如果 $f(x) = 0$ 根是 $\alpha_1, \cdots, \alpha_n$, $g(x) = 0$ 的根是 $\beta_1, \cdots, \beta_m$, 则

$$\Delta = a_0^m b_0^n \prod_{i=1}^{n} \prod_{j=1}^{m} (\alpha_i - \beta_j).$$

这个定理的证明，这儿不谈了.

例 2　算出

$$f(x) = x^3 + px + q = 0$$

有重根的条件. 也便是求出 $f'(x) = 0$, $f(x) = 0$ 有公根的条件.

$$\begin{vmatrix} 1 & 0 & p & q & 0 \\ 0 & 1 & 0 & p & q \\ 3 & 0 & p & 0 & 0 \\ 0 & 3 & 0 & p & 0 \\ 0 & 0 & 3 & 0 & p \end{vmatrix} = \begin{vmatrix} -2p & -3q & 0 \\ 0 & -2p & -3q \\ 3 & 0 & p \end{vmatrix} = 4p^3 + 27q^2.$$

即如果 $f(x)$ 有重根，则 $4p^3 + 27q^2 = 0$.

实质上这个定理不要如此证明的，因为

$$f'(x) = 3x^2 + p.$$

由 $f'(x) = 0$ 解得

$$x = \pm\sqrt{-\frac{p}{3}}.$$

代入 $f(x) = 0$，即得所求.

习题 1　求出

$$x^5 + 5px^3 + 5p^2x + q = 0$$

有重根的条件。

§ 12. 代数曲线的交点

命 $f(x, y)$，$g(x, y)$ 代表 x, y 的多项式；则

$$f(x, y) = 0,$$
$$g(x, y) = 0.$$

各代表一条代数曲线，问题是求这两条代数曲线的交点.

方法是：把 $f(x, y)$ 与 $g(x, y)$ 看成为 x 的多项式，其系数是 y 的函数. 由上节的方法消去 y，即得一个行列式，展开行列式便得 y 的多项式，使行列式=0。解出 y 然后再解出 x 来.

例　求椭圆

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

与双曲线

$$xy = 1$$

的交点。

从

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0,$$
$$yx^2 - x = 0,$$
$$yx - 1 = 0,$$

消去 x^2, x 得

$$\begin{vmatrix} \frac{1}{a_2} & 0 & \frac{y^2}{b^2} - 1 \\ y & -1 & 0 \\ 0 & y & -1 \end{vmatrix} = 0.$$

由此解出 y，然后由 $x=\frac{1}{y}$ 算出 x.

注意　这些方法虽然用行列式的方法表达，但在实际应用时，还是用消去法的好。

§13. 行列式的幂级数

行列式的另一定义是：

$$\begin{vmatrix} a_{22}, & \cdots, & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1}, & \cdots, & a_{nn} \end{vmatrix} = \sum_{i} \delta^{1,\cdots,n}_{i_1,\cdots,i_n} a_{1i_1}\cdots a_{ni_n}.$$

这儿 $\delta_{i_1,\cdots,i_n}=1$，如果 $\begin{pmatrix} 1, & \cdots, & n \\ i_1 & \cdots & i_n \end{pmatrix}$ 是偶置换，$\delta_{i_1,\cdots,i_n}=-1$，如果 $\begin{pmatrix} 1, & \cdots, & n \\ i_1 & \cdots, & i_n \end{pmatrix}$ 是奇置换.

这一定义看来较抽象，并且在具体计算时工作量很为烦重，但是在理论上仍然十分有用，我们现在举一个十分重要的用场.

定理 1　命

$$f_i(z)=\sum_{t=0}^{\infty} a_t^{(i)} z^t,\quad i=1,2\cdots,n,$$

表示 n 个幂级数，并且都当 $|z|<\rho$ 时收敛. 则当

$$|z_1|<\rho,\cdots,|z_n|<\rho$$

时，有次之恒等式

$$\begin{vmatrix} f_1(z_1), & \cdots, & f_1(z_n) \\ \cdots & \cdots & \cdots \\ f_n(z_1), & \cdots, & f_n(z_n) \end{vmatrix} = \sum_{l_1>l_2>\cdots>l_n\geqslant 0} \begin{vmatrix} a_{l_1}^{(1)}, & \cdots, & a_{l_n}^{(1)} \\ \cdots & \cdots & \cdots \\ a_{l_1}^{(n)}, & \cdots, & a_{l_n}^{(n)} \end{vmatrix} \begin{vmatrix} z_1^{l_1}, & \cdots, & z_1^{l_n} \\ \cdots & \cdots & \cdots \\ z_n^{l_1}, & \cdots, & z_n^{l_n} \end{vmatrix}.$$

证明　由行列式的展开法可知此式左边等于

$$\sum_{i_1\cdots i_n} \delta^{1,2\cdots n}_{i_1,i_2,\cdots,i_n} f_{i_1}(z_1)\cdots f_{i_n}(z_n)$$

$$=\sum_{i_1,\cdots,i_n} \delta^{1\cdots n}_{i_1\cdots i_n} \sum_{t_1=0}^{\infty}\cdots\sum_{t_n=0}^{\infty} a_{t_1}^{(i_1)} a_{t_2}^{(i_2)}\cdots a_{t_n}^{(i_n)} z_1^{t_1} z_2^{t_2}\cdots z_n^{t_n}$$

$$=\sum_{t_1=0}^{\infty}\cdots\sum_{t_n=0}^{\infty}\left(\sum_{i_1,\cdots,i_n} \delta^{1,2,\cdots,n}_{i_1,i_2,\cdots,i_n} a_{t_1}^{(i_1)} a_{t_2}^{(i_2)}\cdots a_{t_n}^{(i_n)}\right) z_1^{t_1}\cdots z_n^{t_n}$$

$$=\sum_{t_1=0}^{\infty}\cdots\sum_{t_n=0}^{\infty} \begin{vmatrix} a_{t_1}^{(1)}, & \cdots, & a_{t_n}^{(1)} \\ \cdots & \cdots & \cdots \\ a_{t_1}^{(n)}, & \cdots, & a_{t_n}^{(n)} \end{vmatrix} z_1^{t_1}\cdots z_n^{t_n}$$

$$=\sum_{l_1>l_2>\cdots>l_n\geqslant 0} \begin{vmatrix} a_{l_1}^{(1)} & \cdots & a_{l_n}^{(1)} \\ \cdots & \cdots & \cdots \\ a_{l_1}^{(n)}, & \cdots, & a_{l_n}^{(n)} \end{vmatrix} \sum_{i_1,\cdots,i_n} \delta^{1,2,\cdots,n}_{i_1,\cdots,i_n} z_1^{l_{i_1}}\cdots z_n^{l_{i_n}}.$$

即得本定理.

特别取

$$f_i(z)=f(x_iz)$$

及

$$f(z)=a_0+a_1z+a_2z^2+\cdots,$$

则

$$a_l^{(i)}=a_lx_i^l.$$

代入

$$\begin{vmatrix} a_{l_1}^{(1)}, & \cdots, & a_{l_n}^{(1)} \\ \cdots & \cdots & \cdots \\ a_{l_1}^{(n)}, & \cdots, & a_{l_n}^{(n)} \end{vmatrix}=\begin{vmatrix} a_{l_1}x_1^{l_1}, & \cdots, & a_{l_n}x_1^{l_n} \\ \cdots & \cdots & \cdots \\ a_{l_1}x_n^{l_1}, & \cdots, & a_{l_n}x_n^{l_n} \end{vmatrix}$$

$$=a_{l_1}a_{l_2}\cdots a_{l_n}\begin{vmatrix} x_1^{l_1}, & \cdots, & x_1^{l_n} \\ \cdots & \cdots & \cdots \\ x_n^{l_1}, & \cdots, & x_n^{l_n} \end{vmatrix}.$$

即得

定理 2 命

$$f(z)=a_0+a_1z+\cdots\cdots$$

表示一个当 $|z|<\rho$ 时收敛的幂级数,则当 $|x_iy_j|<\rho$ 时,

$$\begin{vmatrix} f(x_1y_1), & \cdots, & f(x_1y_n) \\ \cdots & \cdots & \cdots \\ f(x_ny_1), & \cdots, & f(x_ny_n) \end{vmatrix}$$

$$=\sum_{l_1>l_2>\cdots>l_n\geqslant 0}a_{l_1}a_{l_2}\cdots a_{l_n}\begin{vmatrix} x_1^{l_1}, & \cdots, & x_n^{l_1} \\ \cdots & \cdots & \cdots \\ x_1^{l_n}, & \cdots, & x_n^{l_n} \end{vmatrix}\begin{vmatrix} y_1^{l_1}, & \cdots, & y_n^{l_1} \\ \cdots & \cdots & \cdots \\ y_1^{l_n}, & \cdots, & y_n^{l_n} \end{vmatrix}.$$

再取特例

$$f(z)=(1-z)^{-1},$$

则有

定理 3 当 $|x_iy_j|<1$ 时,

$$\sum_{l_1>l_2>\cdots>l_n\geqslant 0}\begin{vmatrix} x_1^{l_1}, & \cdots, & x_n^{l_1} \\ \cdots & \cdots & \cdots \\ x_1^{l_n}, & \cdots, & x_n^{l_n} \end{vmatrix}\begin{vmatrix} y_1^{l_1}, & \cdots, & y_n^{l_1} \\ \cdots & \cdots & \cdots \\ y_1^{l_n}, & \cdots, & y_n^{l_n} \end{vmatrix}$$

$$=\frac{\Delta(x_1,\cdots,x_n)\Delta(y_1,\cdots,y_n)}{\prod_{i=1}^{n}\prod_{j=1}^{n}(1-x_iy_j)}.$$

这儿 $\Delta(x_1,\cdots x_n)=\prod_{n\geqslant i>j\geqslant 1}(x_i-x_j)$.

证明 由定理 2 可知,我们需要证明的是

$$\begin{vmatrix} \dfrac{1}{1-x_1y_1}, & \cdots, & \dfrac{1}{1-x_1y_n} \\ \cdots & \cdots & \cdots \\ \dfrac{1}{1-x_ny_1}, & \cdots, & \dfrac{1}{1-x_ny_n} \end{vmatrix} = \frac{\Delta(x_1,\cdots,x_n)\Delta(y_1,\cdots,y_n)}{\prod\limits_{i=1}^{n}\prod\limits_{j=1}^{n}(1-x_iy_j)}.$$

这个恒等式可以从下一定理中换变数即得

定理 4 (Cauchy)

$$\begin{vmatrix} \dfrac{1}{x_1+y_1}, & \cdots, & \dfrac{1}{x_1+y_n} \\ \cdots & \cdots & \cdots \\ \dfrac{1}{x_n+y_1}, & \cdots, & \dfrac{1}{x_n+y_n} \end{vmatrix} = \frac{\Delta(x_1,\cdots,x_n)\Delta(y_1,\cdots,y_n)}{\prod\limits_{i=1}^{n}\prod\limits_{j=1}^{n}(x_i+y_j)}$$

证明 从第二、第三,…第 n 行中各减第一行,且利用

$$\frac{1}{x_l+y_h}-\frac{1}{x_1+y_h}=\frac{x_1-x_l}{(x_1+y_h)(x_l+y_h)},$$
$$l,h=1,2,\cdots,n,$$

可知上面的行列式等于

$$\frac{(x_1-x_2)(x_1-x_3)\cdots(x_1-x_n)}{\prod\limits_{h=1}^{n}(x_1+y_h)}\begin{vmatrix} 1, & \cdots\cdots 1 & \cdots\cdots & 1, \\ \dfrac{1}{x_2+y_1}, & \dfrac{1}{x_2+y_2}, & \cdots, & \dfrac{1}{x_2+y_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{1}{x_n+y_1}, & \dfrac{1}{x_n+y_2}, & \cdots, & \dfrac{1}{x_n+y_n} \end{vmatrix}.$$

由第二第三……第 n 列中各减去第一列,则以上的行列式等于

$$\frac{(y_1-y_2)\cdots(y_1-y_n)}{\prod\limits_{j=2}^{n}(x_j+y_1)}\begin{vmatrix} \dfrac{1}{x_2+y_2}, & \cdots, & \dfrac{1}{x_2+y_n} \\ \cdots & \cdots & \cdots \\ \dfrac{1}{x_n+y_2}, & \cdots, & \dfrac{1}{x_n+y_n} \end{vmatrix}.$$

由归纳法得出本定理.

§14 Wronski 行列式的幂级数展开

由于

$$\begin{vmatrix} z_1^{l_1}, & \cdots, & z_1^{l_n} \\ \cdots & \cdots & \cdots \\ z_n^{l_1}, & \cdots & z_n^{l_n} \end{vmatrix} = \Delta(z_1,\cdots,z_n)P(z_1,\cdots,z_n),$$

此处 P 是 $z_1\cdots z_n$ 的齐次函数,其次数等于

$$l_1+\cdots+l_n-(n-1)-(n-2)\cdots-2-1$$
$$=l_1+\cdots+l_n-\frac{1}{2}n(n-1).$$

由于 $l_1>l_2>\cdots>l_n\geqslant 0$, 所以除

$$l_1 = n-1,\ l_2 = n = 2, \cdots,\ l_{n-1} = 1,\ l_n = 0$$

以外，P 的次数常大于 0，因此当 $z_1 \to 0, \cdots, z_n \to 0$ 时，$P(z_1, z_2, \cdots, z_n) \to 0$. 因此在定理 13.1 的假定下

$$\lim_{\substack{z_1\to 0\\ \cdots\cdots\\ z_n\to 0}} \frac{1}{\Delta(z_1, \cdots, z_n)} \begin{vmatrix} f_1(z_1), & \cdots, & f_1(z_n) \\ \cdots & \cdots & \cdots \\ f_n(z_1), & \cdots, & f_n(z_n) \end{vmatrix}$$

$$= \begin{vmatrix} a_{n-1}^{(1)}, & a_{n-2}^{(1)}, & \cdots & a_1^{(1)}, & a_0^{(1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n-1}^{(n)}, & a_{n-2}^{(n)}, & \cdots, & a_1^{(n)}, & a_0^{(n)} \end{vmatrix}$$

$$= \frac{(-1)^{\frac{1}{2}n(n-1)}}{1!2!\cdots(n-1)!} \begin{vmatrix} f_1(0), & \cdots, & f_n(0) \\ f_1'(0), & \cdots, & f_n'(0) \\ \cdots & \cdots & \cdots \\ f_1^{(n-1)}(0), & \cdots, & f_n^{(n-1)}(0) \end{vmatrix}.$$

换变数可得

定理 1 如果

$$f_1(z), \cdots, f_n(z)$$

在 $z = z_0$ 时解析，则

$$\lim_{\substack{z_1\to z_0\\ \cdots\cdots\\ z_n\to z_0}} \frac{1}{\Delta(z_1, z_2, \cdots, z_n)} \begin{vmatrix} f_1(z_1), & \cdots, & f_1(z_n) \\ \cdots & \cdots & \cdots \\ f_n(z_1), & \cdots, & f_n(z_n) \end{vmatrix}$$

$$= \frac{(-1)^{\frac{1}{2}n(n-1)}}{1!2!\cdots(n-1)!} \begin{vmatrix} f_1(z_0), & \cdots, & f_n(z_0) \\ f_1'(z_0), & \cdots, & f_n'(z_0) \\ \cdots & \cdots & \cdots \\ f_1^{(n-1)}(z_0), & \cdots, & f_n^{(n-1)}(z_0) \end{vmatrix}.$$

习题　用带余项的 Taylor 公式来处理本定理，减弱关于 $f_1(z), \cdots, f_n(z)$ 的限制. 特别取

$$f_i(z) = z^{l_i},\quad l_i > 0,$$

则得

$$\lim_{\substack{z_1\to 1\\ \cdots\cdots\\ z_n\to 1}} \frac{1}{\Delta(z_1, \cdots, z_n)} \begin{vmatrix} z_1^{l_1}, & \cdots, & z_n^{l_1} \\ \cdots & \cdots & \cdots \\ z_1^{l_n}, & \cdots, & z_n^{l_n} \end{vmatrix}$$

$$= \frac{(-1)^{\frac{1}{2}n(n-1)}}{1!2!\cdots(n-1)!} \begin{vmatrix} 1, & \cdots & 1, \\ l_1, & \cdots & l_n, \\ l_1(l_1-1), & \cdots & l_n(l_n-1) \\ \cdots & \cdots & \cdots \end{vmatrix}$$

$$= (-1)^{\frac{1}{2}n(n-1)} \frac{\Delta(l_1, \cdots, l_n)}{1!2!\cdots(n-1)!}.$$

在定理 13.1 中两边除以 $\Delta(z_1, \cdots, z_n)$ 再命 $z_1 \to z, \cdots, z_n \to z$，则得

定理 2 在定理 13.1 假定下

$$\begin{vmatrix} f_1(z), & \cdots, & f_n(z) \\ f_1'(z), & \cdots, & f_n'(z) \\ \cdots & \cdots & \cdots \\ f_1^{(n-1)}(z), & \cdots, & f_n^{(n-1)}(z) \end{vmatrix}$$

$$= \sum_{l_1>l_2>\cdots>l_n\geq 0} \Delta(l_1,\cdots,l_n) \begin{vmatrix} a_{l_1}^{(1)}, & \cdots, & a_{l_n}^{(1)} \\ \cdots & \cdots & \cdots \\ a_{l_1}^{(n)}, & \cdots, & a_{l_n}^{(n)} \end{vmatrix} \times z^{l_1+\cdots+l_n-\frac{1}{2}n(n-1)}$$

此式的左端称为函数 $f_1, \cdots, f_n$ 的 Wronski 行列式.

第二章 矩阵的相抵性

§1. 符 号

我们用 $A=A^{(m,n)}$ 代表 m 行 n 列的矩阵，而 $A=A^{(n)}=A^{(n,n)}$ 表示 n 行列的方阵。单行矩阵 $(m=1)$ 称为矢量(或称为行矢量)，单列矩阵 $(n=1)$ 称为列矢量。如此，$x=x^{(1,m)}$，$b=b^{(1,n)}$，$A=A^{(m,n)}$。

$$xA=b$$

代表线性方程组

$$\sum_{i=1}^{m} x_i a_{ij}=b_j,\quad j=1,2,\cdots,n.$$

对角线方阵以 $[d_1,\cdots,d_n]$ 表之，其中 $d_1,\cdots,d_n$ 顺次是对角线上的元素。

$I(=I^{(n)})$ 表示 n 行列的单位方阵，$0[=0^{(m,n)}]$ 表示 $m\times n$ 的零矩阵。

A' 表示由 A 经行列互换所得出的矩阵。

和与积的定义和一些简单性质都不再重述。用

$$A\begin{pmatrix} s_1, & \cdots, & s_r \\ t_1, & \cdots, & t_r \end{pmatrix}$$

表示矩阵 A 中取第 $s_1,\cdots,s_r$ 行及 $t_1,\cdots,t_r$ 列所做出的 r 行列的方阵的行列式。

定理 1 命

$$A=A^{(m,n)}=(a_{ij}),\ B=B^{(n,l)}=(b_{jk}),$$
$$C=AB\,(=C^{(m,l)}).$$

如果 $p\leqslant n$，则

$$|(c_{ij})_{1\leqslant i,j\leqslant p}|=\sum_{r_1<r_2<\cdots<r_p} A\begin{pmatrix} 1, & 2, & \cdots, & p \\ r_1, & r_2, & \cdots & r_p \end{pmatrix} B\begin{pmatrix} r_1, r_2, & \cdots, & r_p \\ 1, & 2, \cdots, & p \end{pmatrix}.$$

这和过所有的适合于 $r_1<r_2<\cdots<r_p$ 的 r，这 $r_1,\cdots,r_p$ 是从 $1,2,\cdots,n$ 中取来的. 如果 $p>n$，则行列式之值等于 0.

证明 假定 $p\leqslant n$ 由于

$$c_{ij}=\sum_{h=1}^{n} a_{ih}b_{hj},$$

所以

$$|c_{ij}|_{1\leqslant i,j\leqslant p}=\begin{vmatrix} \sum\limits_{s_1=1}^{n} a_{1s_1}b_{s_11}, & \cdots, & \sum\limits_{s_p=1}^{n} a_{1s_p}b_{s_pp} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \sum\limits_{s_1=1}^{n} a_{ps_1}b_{s_11}, & \cdots, & \sum\limits_{s_p=1}^{n} a_{ps_p}b_{s_pp} \end{vmatrix}$$

$$= \sum_{s_1=1}^{n} b_{s_1 1} \begin{vmatrix} a_{1s_1}, \sum_{s_2=1}^{n} a_{1s_2} b_{s_2 2}, & \cdots, & \sum_{s_p=1}^{n} a_{1s_p} b_{s_p p} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{ps_1}, \sum_{s_2=1}^{n} a_{ps_2} b_{s_2 2}, & \cdots, & \sum_{s_p=1}^{n} a_{ps_p} b_{s_p p} \end{vmatrix}$$

$$= \sum_{s_1=1}^{n} \cdots \sum_{s_p=1}^{n} b_{s_1 1} \cdots b_{s_p p} A\begin{pmatrix} 1, 2, & \cdots, & p \\ s_1, s_2, & \cdots, & s_p \end{pmatrix}. \tag{1}$$

此和中有两个 s_h 相等的项一定等于 0，所以我们仅需讨论诸 $s_1, \cdots, s_p$ 互不相等的诸项. 由 $1, 2, \cdots, n$ 中取一组适合于

$$r_1 < r_2 < \cdots < r_p$$

的整数列. 在 $s_1, \cdots, s_p$ 中有 $p!$ 项经过排列可以得这样的数列. 所以

$$|c_{ij}|_{1 \leqslant i, j \leqslant p} = \sum_{r_1 < r_2 \cdots < r_p} \sum_{t_1, \cdots, t_p} A\begin{pmatrix} 1, & 2, \cdots, & p \\ t_1, t_2, & \cdots, & t_p \end{pmatrix} b_{t_1 1} \cdots b_{t_p p}.$$

这儿 $t_1, \cdots, t_p$ 过 $r_1, \cdots, r_p$ 的所有排列. 由于

$$A\begin{pmatrix} 1, & 2, & \cdots, & p \\ t_1, & t_2, & \cdots, & t_p \end{pmatrix} = A\begin{pmatrix} 1, & 2, & \cdots, & p \\ r_1, & r_2, & \cdots, & r_p \end{pmatrix} \delta_{t_1, \cdots, t_p}^{r_1, \cdots, r_p},$$

可知

$$|c_{ij}|_{1 \leqslant i, j \leqslant p} = \sum_{r_1 < r_2 < \cdots < r_p} A\begin{pmatrix} 1, & 2, & \cdots, & p \\ r_1, & r_2, & \cdots, & r_p \end{pmatrix} \sum_{t_1, \cdots, t_p} \delta_{t_1, \cdots, t_p}^{r_1, \cdots, r_p} b_{t_1 1} \cdots b_{t_p p}$$

$$= \sum_{r_1 < r_2 < \cdots < r_p} A\begin{pmatrix} 1, & 2, & \cdots, & p \\ r_1, & r_2, & \cdots, & r_p \end{pmatrix} B\begin{pmatrix} r_1, & \cdots, & r_p \\ 1, & \cdots, & p \end{pmatrix}.$$

即得所证.

如 $p > n$ 则和(1)中所有的项都有两个 s 相等，因此所有的项都是 0，故得定理.

附记 1　同法证明更一般的恒等式

$$C\begin{pmatrix} s_1 & \cdots & s_p \\ t_1, & \cdots & t_p \end{pmatrix} = \sum_{r_1 < \cdots < r_p} A\begin{pmatrix} s_1, & \cdots, & s_p \\ t_1, & \cdots, & r_p \end{pmatrix} B\begin{pmatrix} r_1, & \cdots, & r_p \\ t_1, & \cdots, & t_p \end{pmatrix}.$$

此处 $r_1, \cdots, r_p$ 仍然从 $1, 2, \cdots, n$ 中选取.

附记 2　两个方阵 A, B 之积之行列式等于 A 与 B 的行列式之积.

§2. 秩

如果 A 的 $r+1$ 级子行列式都等于 0 而至少有一个 r 级子行列式 $\neq 0$，则 r 称为 A 的秩.

定理 1　如果 $C = AB$，而 A, B, C 的秩各记为 r_A, r_B, r_C，则

$$r_C \leqslant \min(r_A, r_B).$$

这定理可由上节的结果推出：

如果 A 是可逆方阵，则

$$r_C = r_B.$$

因为由定理 1，可知

$$r_C \leqslant r_B,$$

而又由 $B = A^{-1}C$ 可知

$$r_B \leqslant r_C.$$

同样 B 是可逆方阵，则

$$r_C = r_A.$$

注意 1. 有时取等号，例如

$$A = B = C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

2. 有时取不等号，例如

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad C = 0.$$

定理 2 如果 $C = A + B$，则

$$r_C \leqslant r_A + r_B.$$

证明 取 $p = r_A + r_B + 1$，考虑 C 的任意一个 p 级子行列式

$$\begin{vmatrix} a_{i_1h_1} + b_{i_1h_1}, & \cdots, & a_{i_1h_p} + b_{i_1h_p} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{i_ph_1} + b_{i_ph_1}, & \cdots, & a_{i_ph_p} + b_{i_1h_p} \end{vmatrix}$$

$$= \begin{vmatrix} a_{i_1h_1}, & a_{i_1h_2} + b_{i_1h_2}, & \cdots, & a_{i_1h_p} + b_{i_1h_p} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{i_ph_1}, & a_{i_ph_2} + b_{i_ph_2}, & \cdots, & a_{i_ph_p} + b_{i_ph_p} \end{vmatrix}$$

$$+ \begin{vmatrix} b_{i_1h_1}, & a_{i_1h_2} + b_{i_1h_2}, & \cdots, & a_{i_1h_p} + b_{i_1h_p} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ b_{i_ph_1}, & a_{i_ph_2} + b_{i_ph_2}, & \cdots, & a_{i_ph_p} + b_{i_ph_p} \end{vmatrix}.$$

照这样拆下去，可以把这行列式拆为 2^p 个形如

$$\begin{vmatrix} a_{i_1h_1}, & \cdots, & a_{i_1h_s}, & b_{i_1h_{s+1}}, & \cdots, & b_{i_1h_p} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{i_ph_1}, & \cdots, & a_{i_ph_s}, & b_{i_ph_{s+1}}, & \cdots, & b_{i_ph_p} \end{vmatrix}$$

之和．现在证明这样的行列式 $= 0$，这行列式中或有 r_A 列以上的"a"，或有 r_B 列以上的"b"．即 $s > r_A$ 或 $p - s > r_B$，现在假定 $p - s > r_B$，依第一列展开，展开后再依第二列展开，…，第 s 列展开，最后得如下的形式

$$\sum \pm a_{j_1h_1}a_{j_2h_2}\cdots a_{j_sh_s}\begin{vmatrix} b_{j_{s+1}h_{s+1}}, & \cdots, & b_{j_{s+1}h_p} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ b_{j_ph_{s+1}}, & \cdots, & b_{j_ph_p} \end{vmatrix}.$$

这个和中 $j_1, \cdots, j_p$ 在 $i_1, \cdots, i_p$ 中任取，但 $j_s \neq j_t$，由 $p - s > r_B$ 可知，最后的行列式等于 0，同样处理 $s > r_A$ 的情况．因而得出本定理．

附记 1. $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $r_C < r_A + r_B$.

2. $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $r_C = r_A + r_B$.

§3. 初 等 运 算

在解线性方程组时，我们曾经进行过三种运算：

1. 把一个数 q 乘在第 i 个方程上.

2. 在第 i 个方程上减去第 j 个方程的 t 倍.

3. 把第 i 个方程与第 j 个方程互换.

我们看这三种运算反映在矩阵

$$P=\begin{pmatrix} a_{11}, & \cdots, & a_{1n}, & b_1 \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}, & \cdots, & a_{nn}, & b_n \end{pmatrix}$$

上的意义.

1. 对矩阵的第 i 行上元素同成一数 q 也就是在 P 的左边乘以方阵

$$[1,\ \cdots 1,\ \underset{\text{第 } i \text{ 个}}{q},\ 1,\ \cdots,\ 1],$$

为了方便起见，引进符号 E_{ij}，它是一个 n 行列的方阵，其中除 $a_{ij}=1$ 外，其他的元素都是 0，这样可以写成为

$$[1,\cdots,1,q,1,\cdots,1]=E_{11}+\cdots+E_{i-1i-1}+qE_{ii}+E_{i+1i+1}+\cdots+E_{nn}.$$

2. 每从第 i 行减去第 j 行的对应元素的 t 倍. 也就是在 P 的左边乘以

$$I-tE_{ij}.$$

3. 第 i 行与第 j 行互换，也就是在 P 的左边乘以方阵

$$I-E_{ii}-E_{jj}+E_{ij}+E_{ji}.$$

所以 Gauss 消去法实质上就是在方阵 P 的左边乘以这三类的方阵. 但是第三类方阵是第一、二类方阵的乘积：

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

而

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

第二类的方阵称为平延. 在方阵的研究中有极其重要的地位.

定理 1 任何一个满秩方阵可以表为第一、二类方阵的乘积.

证明 方阵

$$A=(a_{ij})$$

的第一列中至少有一非零的元素，乘以一个第三类方阵把 A 变为一个方阵，其中 $a_{11}\neq 0$，乘以

$$[a_{11}^{-1},\ 1,\cdots,\ 1],$$

不妨假定 $a_{11}=1$，乘以

$$I-a_{21}E_{21}.$$

得一方阵 $a_{21}=0$，依法继行，得 $a_{31}=\cdots=a_{n1}=0$。

再同样考虑

$$\begin{pmatrix} a_{22}, & \cdots, & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{n2}, & \cdots, & a_{2n} \end{pmatrix}.$$

因此可知：一个满秩方阵左边乘上第一、二、三类的方阵可以把它变为三角形

$$\begin{pmatrix} 1 & a_{12} & a_{13}, & \cdots, & a_{1n} \\ 0 & 1, & a_{23}, & \cdots, & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

的形式，再乘以

$$\begin{pmatrix} 1, & -a_{12}, & 0 & \cdots, & 0 \\ 0, & 1, & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

便可把 a_{12} 变为 0 续用此法，最后得出

$$I$$

即得

$$L_1L_2\cdots L_rA=I.$$

此处 L_i 都是第一、二、三类的方阵.

由于这些方阵的逆仍然是这些类型

$([1,1,\cdots,q,1,\cdots,1]^{-1}=[1,\cdots,1,q^{-1},1,\cdots,1];(I-tE_{ij})^{-1}=I+tE_{ij}$，$(I-E_{ii}-E_{jj}+E_{ij}+E_{ji})^{-1}=I-E_{ii}-E_{jj}+E_{ij}+E_{ji})$，故得定理.

附记　这方法可以用来求方阵的逆，仅须记着在进行过程中可以做得"粗"些，例如：

$$\begin{pmatrix} 1, & a_{12}, & \cdots, & a_{1n} \\ a_{21}, & \cdots & \cdots & \cdots \\ \vdots & & & \end{pmatrix}$$

的左边乘上

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -a_{21} & 1 & 0 & \cdots & 0 \\ -a_{31} & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & 0 & & \cdots & 1 \end{pmatrix},$$

可以把第一列第一元素以外的元素一次都化为零. 而这方阵的逆是

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21} & 1 & 0 & \cdots & 0 \\ a_{31} & 0 & 1 & \cdots & \\ \cdots & \cdots & \cdots & \cdots & \end{pmatrix}.$$

定理 2　任一满秩行列式为 1 的方阵一定是平延之积.

证明提要：1)

$$\begin{pmatrix}1 & t\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\s & 1\end{pmatrix}=\begin{pmatrix}1+ts & t\\s & 1\end{pmatrix},$$

$$\begin{pmatrix}1 & 0\\-s(1+ts)^{-1} & 1\end{pmatrix}\begin{pmatrix}1+ts & t\\s & 1\end{pmatrix}=\begin{pmatrix}1+ts & t\\0 & (1+ts)^{-1}\end{pmatrix},$$

$$\begin{pmatrix}1+ts & t\\0 & (1+ts)^{-1}\end{pmatrix}\begin{pmatrix}1 & -(1+ts)^{-1}t\\0 & 1\end{pmatrix}=\begin{pmatrix}1+ts & 0\\0 & (1+ts)^{-1}\end{pmatrix}.$$

因此，$\begin{pmatrix}\lambda & 0\\0 & \lambda^{-1}\end{pmatrix}$ 可以表为平延的乘积.

2）不用 $[a_{11}^{-1},\cdots,1]$，而用 $[a_{11}^{-1}, a_{11}, 1\cdots,1]$.

3）不用 $\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$ 而用 $\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}$.

§4. 相　抵

定义　两个 $m\times n$ 矩阵 A, B 称为相抵，如果有满秩的 $P(=p^{(m)})$, $Q(=q^{(n)})$ 存在使

$$A=PBQ.$$

用符号

$$A\overset{E}{=}B$$

表之.

显然有：1) $A\overset{E}{=}A$; 2)如 $A\overset{E}{=}B$,则 $B\overset{E}{=}A$; 3) 如 $A\overset{E}{=}B$, $B\overset{E}{=}C$, 则 $A\overset{E}{=}C$.

定理 1　相抵的必要且充分条件是 A 与 B 的秩相等. 凡秩等于 r 的矩阵一定相抵于

$$\begin{pmatrix}I^{(r)} & 0\\0 & 0\end{pmatrix}.$$

证明　经过换行换列(即左乘及右乘以第三类的方阵)不仿假定

$$A=\begin{pmatrix}A_1 & A_2\\A_3 & A_4\end{pmatrix},\ A_1=A_1^{(r)},\ (A_1)\neq 0,$$

左乘以

$$P=\begin{pmatrix}A_1^{-1} & 0\\-A_3A_1^{-1} & I\end{pmatrix},$$

得

$$\begin{pmatrix}I & A_1^{-1}A_2\\0 & *\end{pmatrix}.$$

这矩阵的秩$=r$，所以"*"处等于0，取

$$Q=\begin{pmatrix}I, & -A_1^{-1}A_2\\0, & I\end{pmatrix}$$

即得所求.

秩等于 1 的矩阵可以表为

$$\begin{pmatrix} a_1b_1, & a_1b_2, & \cdots, & a_1b_n \\ \cdots & \cdots & \cdots & \cdots \\ a_mb_1, & a_mb_2, & \cdots, & a_mb_n \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} (b_1, \cdots, b_n).$$

读者试由定理 1 来推出此结果.

如果A的秩等于m，则可以凑上$n-m$行使

$$\begin{pmatrix} A \\ C \end{pmatrix}$$

是一个可逆方阵。

证明　由于

$$A = P(I, 0)Q,$$

添上

$$C = (0, I)Q,$$

而

$$\begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} P & 0 \\ 0 & I \end{pmatrix} Q.$$

即得所求。

§5. n维矢量空间

R_n 代表所有的矢量

$$x = (x_1, \cdots, x_n)$$

的集合，称为n维矢量空间.

我们不再复习线性相关及线性无关的一些基本性质，我们只强调一点．m个矢量

$$x^{(j)} = (x_1^{(j)}, \cdots, x_n^{(j)}), \quad j = 1, 2, \cdots, m$$

线性相关与否取决于矩阵

$$X = \begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(m)} \end{pmatrix} = \begin{pmatrix} x_1^{(1)}, & \cdots, & x_n^{(1)} \\ \cdots & \cdots & \cdots \\ x_1^{(m)}, & \cdots, & x_n^{(m)} \end{pmatrix}$$

的秩是否等于m. 即X的秩是m，则 $x^{(1)}, \cdots, x^{(m)}$ 线性无关．不然则线性相关．假定其秩等于r，则其中有r个线性无关的矢量，其他的可以表为这几个矢量的线性组合.

假定

$$x^{(1)}, \cdots, x^{(m)}$$

是m个线性无关的矢量，则由

$$x = c_1x^{(1)} + \cdots + c_mx^{(m)}$$

所表出的矢量的全体称为m维子空间．这m个矢量称为张开这子空间的基本矢量，X称为这个子空间的表达矩阵．如果另有l个矢量

$$y^{(1)}, \cdots, y^{(l)}$$

也张开这子空间．则

$$x^{(i)}=\sum_{j=1}^{l}c_{ij}y^{(j)},\quad 1\leqslant i\leqslant m.$$

即

$$X=CY.$$

由此得 $r(X)\leqslant r(Y)$，同样 $r(Y)\leqslant r(X)$，因此 $r(X)=r(Y)$，因而 $l=m$，即得

X，Y 代表同一子空间，则有一满秩方阵 Q 使

$$X=QY. \tag{1}$$

例 $(1,0,\cdots,0)$ 代表一个一维子空间，其中的所有的元素是 $(x,0,\cdots,0)$，对固定的 $a_1,\cdots,a_n$。$(a_1t,\cdots,a_nt)$ 也代表一个一维子空间.

关系 (1) 称为"左相抵". 显然也是一个等价关系.

§6. 矢量空间的变换

变换

$$y=xA,\ A=A^{(n)} \tag{1}$$

把 R_n 中一个矢量 x 变为一个矢量 y.

如果 A 是满秩的，则可以解得

$$x=yA^{-1}. \tag{2}$$

这称为 (1) 的逆变换. 这样的变换 (2) 把 R_n 一一对应地变为其自己.

继行

$$z=yB \tag{3}$$

得

$$z=x(AB).$$

即连续两次变换的方阵等于两个方阵之积.

在 R_n 中有 m 个矢量

$$x^{(1)},\cdots,x^{(m)},$$

则经 (1) 变为 m 个矢量

$$y^{(i)}=x^{(i)}A.$$

写成矩阵形式

$$Y=XA.$$

"相抵"关系的几何意义是 X 所代表的 m 维子空间，经变形(1)变为

$$Y=QXA$$

所代表的 m 维子空间.

由相抵定理立刻推得

定理 1 任何一个 m 维线性子空间，可以由 (1) 变为由

$$\varepsilon_1=(1,0,\cdots,0),\quad \varepsilon_2=(0,1,0,\cdots0)\cdots,\quad \varepsilon_m=(0,\cdots,1,0\cdots0)$$

所张成的子空间，也可以述为：在变换"群" (1) 之下，m 维子空间成一"可递"集合.

特别有：任一矢量可以由 (1) 变为 $(1,0,\cdots,0)$，或对任意两个矢量我们有一个 (1) 把他们变来变去.

我们不假定 A 是可逆的．假定 A 的秩等于 r，则 (1) 把 R_n 映射成为一 r 维子空间．原因是：

$$PAQ=\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix},\quad |P|\neq 0,\ |Q|\neq 0.$$

而

$$y = xA$$

变为

$$yQ = xP^{-1}\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

命 $y^*=yQ$，$x^*=xP^{-1}$，则 $y^*=(\underbrace{*\cdots*}_{r\text{个}},0,\cdots,0)$，这些 y^* 成一 r 维子空间．因此所有的 y 也成一 r 维子空间．

其次考虑 (1) 把那些 x 变为 0，求解

$$0 = xA,$$

即

$$0 = xP^{-1}\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

因而得

$$x^*=(\overbrace{0,\cdots,0}^{r\text{个}},*,\cdots,*).$$

所以 (1) 把一个 $n-r$ 维的子空间映射为零．

§7. 长度、角度与面积等

现在考虑实矢量．

定义　两个矢量 a,b 的内积定义为

$$ab' = \sum_{i=1}^{n} a_i b_i = ba'.$$

一个矢量的长度 $=\sqrt{aa'}=\sqrt{a_1^2+\cdots+a_n^2}$．两个矢量 a、b 所夹成的三角形的三边的长度是

$$\sqrt{aa'},\ \sqrt{bb'},\ \sqrt{(a-b)(a-b)'},$$

由余弦定理得

$$(a-b)(a-b)' = aa' + bb' - 2\sqrt{aa'bb'}\cos\theta, \tag{1}$$

这儿 θ 是矢量 a 和矢量 b 的夹角．由 (1) 可知

$$\cos\theta = \frac{ab'}{\sqrt{aa'bb'}}.$$

$ab'=0$，这二矢量定义为正交，或相互垂直．

这三角形的面积等于

$$\Delta = \frac{1}{2}\sqrt{aa'bb'}\sin\theta = \frac{1}{2}\sqrt{aa'bb'-(ab')^2}.$$

平方根号下的数等于方阵

$$\begin{pmatrix} \underline{a}\underline{a}', & \underline{a}\underline{b}' \\ \underline{b}\underline{a}', & \underline{b}\underline{b}' \end{pmatrix} = \begin{pmatrix} \underline{a} \\ \underline{b} \end{pmatrix} \begin{pmatrix} \underline{a} \\ \underline{b} \end{pmatrix}$$

的行列式的平方根，由于

$$\underline{a}\underline{a}'\underline{b}\underline{b}' - (\underline{a}\underline{b}')^2 = \sum_{1 \leqslant i < j \leqslant n} (a_i b_j - a_j b_i)^2,$$

这等于 $\begin{pmatrix} \underline{a} \\ \underline{b} \end{pmatrix}$ 中所有的二行列的子行列式的平方和.

一般讲来(我们现在不证)：m个矢量

$$\underline{a}^{(1)}, \cdots, \underline{a}^{(m)}$$

所形成的"m 维单纯形"所谓单纯形：0 及 $\underline{a}^{(i)}(i = 1, \cdots, m)$ 的终点称为顶点过其中的 m点有一 $(m-1)$ 维平面，$m+1$ 个平面所围的体称为m维单纯形作出矩阵

$$A = \begin{pmatrix} \underline{a}^{(1)} \\ \vdots \\ \underline{a}^{(m)} \end{pmatrix}.$$

则这个单纯形的体积等于

$$\frac{1}{m!}|AA'|^{\frac{1}{2}}.$$

最简单的单纯形的例子是由 $\underline{e}_1 = (1, 0, \cdots, 0)$, $\underline{e}_2 = (0,1,0,\cdots,0), \cdots$, $\underline{e}_n = (0, 0, \cdots, 1)$ 所形成的单纯形，也就是

$$x_1 \geqslant 0, \cdots, x_n \geqslant 0, \quad x_1 + \cdots + x_n \leqslant 1.$$

其次

$$0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n \leqslant 1.$$

也是一个单纯形，命 $t_1 = x_1$, $t_2 = x_1 + x_2, \cdots$, $t_n = x_1 + \cdots + x_n$ 则第二个形立刻变为第一个形，第二个形的体积是单位立方体的 $\frac{1}{n!}$. 这是显然的，因为 $0 \leqslant x_i \leqslant 1(i = 1, \cdots, n)$ 可以分为 $n!$ 块，而每一块经过变数排列就可以变为以上的形式.

研究体积时还必须考虑序向. 例如：在平面上反时针方向为正，顺向为负，在三维空间如果三矢量的次数成左手座标为正，右手为负，在 n 维空间如果我们固定了次序

$$\underline{e}_1, \underline{e}_2, \cdots, \underline{e}_n.$$

则经过次序偶排列的为正，奇排列的为负，对 n 个矢量

$$\underline{a}^{(1)}, \cdots, \underline{a}^{(n)}$$

来说，行列式

$$\begin{vmatrix} \underline{a}^{(1)} \\ \vdots \\ \underline{a}^{(n)} \end{vmatrix}.$$

为正的是正序向，为负的是负序向.

§8. 函数行列式 (Jacobian)

已往考虑过线性变换

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \cdots, m.$$

我们现在考虑非常一般的变换

$$y_i = \varphi_i(x_1, \cdots, x_n), \quad i = 1, 2, \cdots, m. \tag{1}$$

假定这些 $\varphi_1, \cdots, \varphi_m$ 都是可以求偏微商的，如此，则

$$dy_i = \sum_{j=1}^{n} dx_j \frac{\partial \varphi_i}{\partial x_j} = \sum_{j=1}^{n} dx_j \frac{\partial y_i}{\partial x_j}, \quad i = 1, 2, \cdots, m.$$

写成矩阵符号得(在不致于引起误解时，用 x 代替 $\underline{x}$ 表示矢量)

$$dy = dx \frac{\partial(y_1, \cdots, y_m)}{\partial(x_1, \cdots, x_n)},$$

这儿 $dx = (dx_1, \cdots, dx_n)$ 是微分矢量，而矩阵

$$\frac{\partial(y_1, \cdots, y_m)}{\partial(x_1, \cdots, x_n)} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1}, & \cdots, & \frac{\partial y_m}{\partial x_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_1}{\partial x_n}, & \cdots, & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

称 Jacobian 矩阵. 如果 $z_1, \cdots, z_l$ 又是 $y_1, \cdots, y_m$ 的函数，则

$$dz = dy \frac{\partial(z_1, \cdots, z_l)}{\partial(y_1, \cdots, y_m)}.$$

因此

$$dz = dx \frac{\partial(y_1, \cdots, y_m)}{\partial(x_1, \cdots, x_n)} \frac{\partial(z_1, \cdots, z_l)}{\partial(y_1, \cdots, y_m)}.$$

即得 Jacobian 矩阵的乘积法则

$$\frac{\partial(z_1, \cdots, z_l)}{\partial(x_1, \cdots, x_n)} = \frac{\partial(y_1, \cdots, y_m)}{\partial(x_1, \cdots, y_n)} \frac{\partial(z_1, \cdots, z_l)}{\partial(y_1, \cdots, y_m)}.$$

当 $m = n = l$ 时，如果 $z_1 = x_1, \cdots, z_n = x_n$，则得

$$I = \frac{\partial(y_1, \cdots, y_m)}{\partial(x_1, \cdots, x_m)} \frac{\partial(x_1, \cdots, x_m)}{\partial(y_1, \cdots, y_m)}.$$

即如果变换(1)的逆变换存在，则逆变换的 Jacobian 方阵等于原变换 Jacobian 方阵之逆.

§9. 隐函数定理

定理 1 假定方程组

$$F_1(x_1, \cdots, x_m, y_1, \cdots, y_n) = 0, \cdots, F_n(x_1, \cdots, x_m, y_1, \cdots, y_n) = 0 \tag{1}$$

有一个解

$$x_k = x_k^{(0)}, \ y_l = y_l^{(0)}, \ 1 \leqslant k \leqslant m, \ 1 \leqslant l \leqslant n. \tag{2}$$

假定 F_l 在值 (2) 的数域内为连续函数，而且有一级偏微商，并且假定在值 (2) 时

$$\frac{\partial(F_1, \cdots F_n)}{\partial(y_1, \cdots, y_n)} \neq 0.$$

于是，当 x_k 充分接近于 $x_k^{(0)}$ 时，方程组(1)确定一函数组 $y_l(x_1, \cdots, x_m)$，它们是连续的，

具有一级微商而且满足条件

$$y_l(x_1^{(0)},\cdots,x_m^{(0)})=y_l^{(0)}.$$

证明　对 n 作归纳法，当 $n=1$ 时已知此定理正确，假定对 $(n-1)$ 个方程此定理成立，今后证对 n 个方程也成立．

经过适当地改换函数的号码，可以假定在值 (2) 处

$$\frac{\partial(F_2,\cdots,F_n)}{\partial(y_2,\cdots,y_n)}\neq 0.$$

因此唯一地确定函数组

$$y_2=\varphi_2(x_1,\cdots,x_m,y_1),\cdots,\quad y_n=\varphi_n(x_1,\cdots,x_m,y_1).\tag{3}$$

以此代入 F_1 中得

$$F_1(x_1,\cdots,x_m,y_1,\varphi_2,\cdots,\varphi_n)=0.\tag{4}$$

这是 $x_1,\cdots,x_m,y_1$ 的函数，如果能证明它对 y_1 的偏微商不等于 0，则可由此得出 y_1，y_1 是 $x_1,\cdots,x_m$ 的适合于定理所要求的函数，再代入 (3) 即得定理．

因此问题变为求证(4)对 y_1 的微商(指把 $\varphi_2,\cdots,\varphi_n$ 代入得的情况)

$$\left(\frac{\partial F_1}{\partial y_1}\right)=\frac{\partial F_1}{\partial y_1}+\sum_{s=2}^{n}\frac{\partial F_1}{\partial\varphi_s}\frac{\partial\varphi_s}{\partial y_1}\neq 0,\tag{5}$$

再由

$$F_l(x_1,\cdots,x_m,y_1,\varphi_2,\cdots,\varphi_n)=0$$

可知

$$\frac{\partial F_l}{\partial y_1}+\sum_{s=2}^{n}\frac{\partial F_l}{\partial\varphi_s}\frac{\partial\varphi_s}{\partial y_1}=0,\quad l=2,\cdots,n.\tag{6}$$

在(5),(6)中消去 $1,\dfrac{\partial\varphi_2}{\partial y_1},\cdots,\dfrac{\partial\varphi_n}{\partial y_1}$ 得

$$\begin{vmatrix}-\left(\dfrac{\partial F_1}{\partial y_1}\right)+\dfrac{\partial F_1}{\partial y_1}, & \dfrac{\partial F_1}{\partial\varphi_2}, & \cdots, & \dfrac{\partial F_1}{\partial\varphi_n}\\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots\\ \dfrac{\partial F_n}{\partial y_1}, & \dfrac{\partial F_n}{\partial\varphi_2}, & \cdots, & \dfrac{\partial F_n}{\partial\varphi_n}\end{vmatrix}=0.$$

即得

$$\left(\frac{\partial F_1}{\partial y_1}\right)\frac{\partial(F_2,\cdots,F_l)}{\partial(y_2,\cdots,y_n)}-\frac{\partial(F_1,\cdots,F_n)}{\partial(y_1,\cdots,x_n)}=0.$$

所以

$$\left(\frac{\partial F_1}{\partial y_1}\right)\neq 0.$$

由此定理立刻推出反演定理．

定理 2　设有

$$y_i=f_i(x_1,\cdots,x_n),\quad i=1,2,\cdots,n.\tag{7}$$

假定 f_i 以及其一级偏微商在 $x_i=x_i^{(0)}$ $(i=1,\cdots,n)$ 的邻域内连续而且在此点

$$\frac{\partial(y_1,\cdots,y_n)}{\partial(x_1,\cdots,x_n)}\neq 0,$$

于是方程组 (7) 唯一地确定一组在值

$$y_i^{(0)} = f_i(x_1^{(0)}, \cdots, x_n^{(0)})$$

的邻域作为 $y_1, \cdots, y_n$ 的函数看的 $x_k(y_1, \cdots, y_n)$，这些函数是连续的，具有一级微商的，并且适合于

$$x_k^{(0)} = x_k(y_1^{(0)}, \cdots, y_n^{(0)}).$$

这证明是上述定理取

$$F_i = f_i - y_i$$

的特例.

§10. 复变函数的 Jacobian

如果

$$f_j = \sum_{k=1}^{n} \varphi_k c_{kj}, \quad j = 1, 2, \cdots, n.$$

$$x_j = \sum_{k=1}^{n} y_k c_{kj},$$

则

$$\frac{\partial(\varphi_1, \cdots, \varphi_n)}{\partial(y_1, \cdots, y_n)} = \frac{\partial(x_1, \cdots, x_n)}{\partial(y_1, \cdots, y_n)} \frac{\partial(f_1, \cdots, f_n)}{\partial(x_1, \cdots, x_n)} \frac{\partial(\varphi_1, \cdots, \varphi_n)}{\partial(f_1, \cdots, f_n)}$$
$$= C \frac{\partial(f_1, \cdots, f_n)}{\partial(x_1, \cdots, x_n)} C^{-1},$$

因此

$$\det\left(\frac{\partial(\varphi_1, \cdots, \varphi_n)}{\partial(y_1, \cdots, y_n)}\right) = \det\left(\frac{\partial(f_1, \cdots, f_n)}{\partial(x_1, \cdots, x_n)}\right).$$

根据这个结果，我们来证明.

定理 1 假定

$$f_j(z_1, \cdots, z_n), \quad j = 1, 2, \cdots, n$$

是 n 个复变数函数的解析函数，把变数和函数都分为虚实部分

$$z_k = x_k + iy_k, \quad f_k = u_k + iv_k,$$

则得

$$\det\left(\frac{\partial(u_1, v_1, \cdots, u_n, v_n)}{\partial(x_1, y_1, \cdots, x_n, y_n)}\right) = \left|\det \frac{\partial(f_1, \cdots, f_n)}{\partial(z_1, \cdots, z_n)}\right|^2.$$

证明

$$(z_j, \bar{z}_j) = (x_j, y_j)\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix},$$

$$(f_j, \bar{f}_j) = (u_j, v_j)\begin{pmatrix} 1, & 1 \\ i, & -i \end{pmatrix}.$$

因此

$$\det \frac{\partial(u_1, v_1, \cdots, u_n, v_n)}{\partial(x_1, y_1, \cdots, x_n, y_n)} = \det\left(\frac{\partial(f_1, \bar{f}_1, \cdots, f_n, \bar{f}_n)}{\partial(z_1, \bar{z}_1, \cdots, z_n, \bar{z}_n)}\right)$$
$$= \det \frac{\partial(f_1, \cdots, f_n; \bar{f}_1, \bar{f}_2, \cdots, \bar{f}_n)}{\partial(z_1, \cdots, z_n; \bar{z}_1, \cdots, \bar{z}_n)}.$$

如果 $f(z)$ 是 z 的解析函数，则由定义可知 $\frac{\partial f(z)}{\partial \bar{z}}=0$，$\frac{\partial \bar{f}}{\partial z}=0$。因此上一行列式等于

$$\det \frac{\partial(f_1,\cdots,f_n)}{\partial(z_1,\cdots,z_n)} \det \frac{\partial(\bar{f}_1,\cdots,\bar{f}_n)}{\partial(\bar{z}_1,\cdots,\bar{z}_n)}$$

即得所证.

关于解析函数的隐函数定理有:

定理 2 如果在原点附近 n 个函数

$$F_j(w_1,\cdots,w_n;\ z_1,\cdots,z_n),\quad j=1,\cdots,n$$

是解析的，而且

$$F_j(0,\cdots,0)=0$$

及

$$\det\left(\frac{\partial(F_1,\cdots,F_n)}{\partial(w_1,\cdots,w_n)}\right)\neq 0 \quad w=0,\ z=0.$$

则方程组

$$F_j(w_1,\cdots,w_n,\ z_1,\cdots,z_n)=0,\quad j=1,\cdots,n \tag{1}$$

有唯一解

$$w_j=w_j(z_1,\cdots,z_n),\quad j=1,2,\cdots,n, \tag{2}$$

此解当 $z=0$ 时，$w=0$ 并且在 $z=0$ 附近是解析的.

证明　我们把方程(1)分为虚实部分，则得出 $2n$ 个方程式，$w_j=u_j+iv_j$ 看成为 $2n$ 个实变数 u_i，v_i，由定理 1，可知作为实变数来说，Jacobian 在原点非零，因此可以解出一个解有一阶的连续偏微商，这解是唯一的，而且在原点处=0。

对 $\bar{z}_m$ 微分(1)式得

$$\sum_{p=1}^{n}\frac{\partial F_j}{\partial w_p}\frac{\partial w_p}{\partial \bar{z}_m}+\sum_{\lambda=1}^{l}\frac{\partial F_j}{\partial z_\lambda}\frac{\partial z_\lambda}{\partial \bar{z}_m}=0.$$

由于 $\partial z_\lambda/\partial \bar{z}_m=0$，所以

$$\sum_{p=1}^{n}\frac{\partial F_j}{\partial w_p}\frac{\partial w_p}{\partial \bar{z}_m}=0,\quad j,m=1,\cdots,n.$$

由于 $\det\left(\frac{\partial(F_1,\cdots,F_n)}{\partial(w_1,\cdots,w_n)}\right)\neq 0$，所以 $\frac{\partial w_p}{\partial \bar{z}_m}=0$，即 w_p 是 $z_1,\cdots,z_n$ 的解析函数.

§11. 函 数 相 关

现在考虑 n 个变数的 m 个函数

$$\left.\begin{array}{l} u_1=u_1(x_1,\cdots,x_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ u_m=u_m(x_1,\cdots,x_n). \end{array}\right\} \tag{1}$$

假定这些函数是连续的，而且有一阶连续偏微商，如果有一个非零函数

$$F(u_1,\cdots,u_m),$$

当把(1)代入此函数时，得到一个对 $x_1,\cdots,x_n$ 恒等于 0 的式子，这 m 个函数称为函数相关.

我们主要的结果是：考虑 Jacobian 矩阵

$$\frac{\partial(u_1,\cdots,u_m)}{\partial(x_1,\cdots,x_n)}=\begin{pmatrix}\frac{\partial u_1}{\partial x_1}, & \cdots, & \frac{\partial u_1}{\partial x_n}\\ \cdots & \cdots & \cdots\\ \frac{\partial u_m}{\partial x_1}, & \cdots, & \frac{\partial u_m}{\partial x_n}\end{pmatrix}.$$

如果它所有的 $r+1$ 行列的子行列式都恒等于 0，而有一个 r 行列的子行列式不恒等于 0，则有 $m-r$ 个可以表为其他的 u 的函数。当 $n=m=r$ 时，$u_1,\cdots,u_m$ 非函数互依。

如果 $m<n$，我们不妨添 $n-m$ 个函数

$$u_{m+1}=x_{m+1},\cdots,u_n=x_n.$$

如果 $m>n$，不妨假定函数是与另加的 $m-n$ 个变数 $x_{n+1},\cdots,x_m$ 无关的函数，因此今后不妨假定 $m=n$。

这一结果稍为复杂些，我们现在分步分段的说明如下（以下是分析学者的处理方法，我们可以看出逐步思考的过程，下节中我们将归结为代数处理方法）。

为了不与 Jacobian 矩阵 相混淆，我们用

$$\frac{D(u_1,\cdots,u_n)}{D(x_1,\cdots,x_n)}$$

表示 Jacobian 方阵 $\frac{\partial(u_1,\cdots,u_n)}{\partial(x_1,\cdots,x_n)}$ 的行列式，或简称为 Jacobian。

定理 1 命 $m=n$，(1) 所定义的函数为函数相关的必要且充分条件是

$$\frac{D(u_1,\cdots,u_n)}{D(x_1,\cdots,x_n)}\equiv 0. \tag{2}$$

证明

1）必要性 为了容易理解，我们取 $n=3$：

$$X=f_1(x,y,z),\quad Y=f_2(x,y,z),\quad Z=f_3(x,y,z). \tag{3}$$

假定

$$\frac{D(X,Y,Z)}{D(x,y,z)}\not\equiv 0.$$

即有一点 $x=x_0$，$y=y_0$，$z=z_0$（由 (1) 得出对应点 $X=X_0$，$Y=Y_0$，$Z=Z_0$）使

$$\left(\frac{D(X,Y,Z)}{D(x,y,z)}\right)_{x=x_0,y=y_0,z=z_0}\neq 0.$$

由隐函数定理有 h 存在使适合于

$$|X-X_0|\leqslant h,\ |Y-Y_0|\leqslant h,\ |Z-Z_0|\leqslant h \tag{4}$$

的每一个 (X,Y,Z) 都有 x,y,z 适合于 (1)，换言之，f_1,f_2,f_3 能取区间 (4) 中的任意数值，因此不可能有函数关系

$$F(X,Y,Z)=0$$

存在。

同样理由可以证明，如果

$$\frac{D(X,Y)}{D(x,y)},\ \frac{D(X,Y)}{D(y,z)},\ \frac{D(X,Y)}{D(z,x)}$$

中有一个不恒等于 0，则 X，Y 间没有函数关系。

更一般些，如果 $n>m$，而(1)所定义的m个函数 $u_1,\cdots,u_m$ 有函数关系则

$$\frac{D(u_1,\cdots,u_m)}{D(x_{\alpha_1},\cdots,x_{\alpha_m})}=0.$$

此处 $\alpha_1,\cdots,\alpha_m$ 是取自 $1,\cdots,n$ 中的任意m个，也就是 Jacobian 矩阵的“秩”(指恒等)应当 $<m$.

2）充分性　取四个变数为例

$$\left.\begin{aligned}X&=f_1(x,y,z,t)\\Y&=f_2(x,y,z,t)\\Z&=f_3(x,y,z,t)\\T&=f_4(x,y,z,t)\end{aligned}\right\}\tag{5}$$

并且假定

$$\Delta=\begin{vmatrix}\frac{\partial f_1}{\partial x}&\frac{\partial f_1}{\partial y}&\frac{\partial f_1}{\partial z}&\frac{\partial f_1}{\partial t}\\\frac{\partial f_2}{\partial x}&\frac{\partial f_2}{\partial y}&\frac{\partial f_2}{\partial z}&\frac{\partial f_2}{\partial t}\\\frac{\partial f_3}{\partial x}&\frac{\partial f_3}{\partial y}&\frac{\partial f_3}{\partial z}&\frac{\partial f_3}{\partial t}\\\frac{\partial f_4}{\partial x}&\frac{\partial f_4}{\partial y}&\frac{\partial f_4}{\partial z}&\frac{\partial f_4}{\partial t}\end{vmatrix}=0.$$

先假定

$$\delta=\frac{D(f_1,f_2,f_3)}{D(x,y,z)}\neq 0,$$

则可解得

$$x=\varphi_1(X,Y,Z,t),\quad y=\varphi_2(X,Y,Z,t),\quad z=\varphi_3(X,Y,Z,t).\tag{6}$$

代入(5)中第四式得

$$T=f_4(\varphi_1,\varphi_2,\varphi_3,t)=F(X,Y,Z,t).$$

我们只须证明F与t无关，即

$$\frac{\partial F}{\partial t}=0.$$

由于

$$\frac{\partial F}{\partial t}=\frac{\partial f_4}{\partial x}\frac{\partial \varphi_1}{\partial t}+\frac{\partial f_4}{\partial y}\frac{\partial \varphi_2}{\partial t}+\frac{\partial f_4}{\partial z}\frac{\partial \varphi_3}{\partial t}+\frac{\partial f_4}{\partial t},\tag{7}$$

这儿 $\frac{\partial\varphi_1}{\partial t}$，$\frac{\partial\varphi_2}{\partial t}$，$\frac{\partial\varphi_3}{\partial t}$ 是隐函数 $\varphi_1,\varphi_2,\varphi_3$ 对 t 的微商，由(5)来决定 $\frac{\partial\varphi_1}{\partial t}$，$\frac{\partial\varphi_2}{\partial t}$，$\frac{\partial\varphi_3}{\partial t}$.

$$\left.\begin{aligned}\frac{\partial f_1}{\partial x}\frac{\partial \varphi_1}{\partial t}+\frac{\partial f_1}{\partial y}\frac{\partial \varphi_2}{\partial t}+\frac{\partial f_1}{\partial z}\frac{\partial \varphi_3}{\partial t}+\frac{\partial f_1}{\partial t}=0\\\frac{\partial f_2}{\partial x}\frac{\partial \varphi_1}{\partial t}+\frac{\partial f_2}{\partial y}\frac{\partial \varphi_2}{\partial t}+\frac{\partial f_2}{\partial z}\frac{\partial \varphi_3}{\partial t}+\frac{\partial f_2}{\partial t}=0\\\frac{\partial f_3}{\partial x}\frac{\partial \varphi_1}{\partial t}+\frac{\partial f_3}{\partial y}\frac{\partial \varphi_2}{\partial t}+\frac{\partial f_3}{\partial z}\frac{\partial \varphi_3}{\partial t}+\frac{\partial f_3}{\partial t}=0.\end{aligned}\right\}\tag{8}$$

由(7)与(8)消去 $\frac{\partial\varphi_1}{\partial t}$, $\frac{\partial\varphi_2}{\partial t}$, $\frac{\partial\varphi_3}{\partial t}$, 1 可得

$$\begin{vmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} & & \frac{\partial f_1}{\partial t} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} & & \frac{\partial f_2}{\partial t} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} & & \frac{\partial f_3}{\partial t} \\ \frac{\partial f_4}{\partial x} & \frac{\partial f_4}{\partial y} & \frac{\partial f_4}{\partial z} & \frac{\partial f_4}{\partial t} & -\frac{\partial F}{\partial t} \end{vmatrix} = 0.$$

即得

$$0 = \Delta - \delta\frac{\partial F}{\partial t}.$$

由于 $\delta \neq 0$, 所以 $\frac{\partial F}{\partial t} = 0$, 即函数 X, Y, Z, T 间有关系

$$T = F(X, Y, Z). \tag{9}$$

注意　除这个函数关系之外,不存在其他关系,如不然,以(9)代人,得一个 X, Y, Z 之间的关系,因而 $\delta = 0$, 这与假定不符.

现在再考虑△的所有的三阶子行列式都$=0$的情况，假定有一个二阶子行列式

$$\delta' = \frac{D(f_1, f_2)}{D(x, y)} \neq 0.$$

从(5)的前二式解得

$$x = \varphi_1(X, Y, z, t), \qquad y = \varphi_2(X, Y, z, t).$$

因此

$$Z = f_3(x, y, z, t) = F_1(X, Y, z, t),$$
$$T = f_4(x, y, z, t) = F_2(X, Y, z, t).$$

由

$$\frac{\partial F_1}{\partial t} = \frac{\partial f_3}{\partial x}\frac{\partial\varphi_1}{\partial t} + \frac{\partial f_3}{\partial y}\frac{\partial\varphi_2}{\partial t} + \frac{\partial f_3}{\partial t},$$
$$0 = \frac{\partial f_1}{\partial x}\frac{\partial\varphi_1}{\partial t} + \frac{\partial f_1}{\partial y}\frac{\partial\varphi_2}{\partial t} + \frac{\partial f_1}{\partial t},$$
$$0 = \frac{\partial f_2}{\partial x}\frac{\partial\varphi_1}{\partial t} + \frac{\partial f_2}{\partial y}\frac{\partial\varphi_2}{\partial t} + \frac{\partial f_2}{\partial t},$$

推得

$$\frac{D(f_1, f_2, f_3)}{D(x, y, z)} = \delta'\frac{\partial F_1}{\partial t},$$

因此 $\frac{\partial F_1}{\partial t} = 0$. 同法证明 $\frac{\partial F_1}{\partial z} = 0$, 因此 F_1 中不会有 t 与 z, 同样处理, F_2 也与 t, z 无关,即有两个关系

$$Z = F_1(X, Y), \qquad T = F_2(X, Y).$$

最后再处理所有的二阶子行列式都 $= 0$, 设有一阶子行列式$\neq 0$, 这样 X, Y, Z, T

中有三个能表为另一个函数.

总之：如果

$$\frac{\partial(F_1,\cdots,F_n)}{\partial(x_1,\cdots,x_n)}$$

的秩等于 r，则 $F_1,\cdots,F_n$ 中有 $n-r$ 个可以表为其他 r 个的函数，即函数无关的函数的个数等于 r. 而且没有其他的函数关系.

因此不难推得本节的开始的结论.

特例：两个函数 $F_1(x_1,\cdots,x_n)$，$F_2(x_1,\cdots,x_n)$ 是彼此相关的函数的条件是

$$\frac{\partial F_1}{\partial x_i}:\frac{\partial F_2}{\partial x_i}$$

与 i 无关.

附记 1　若以上所说的函数

$$F_1, F_2, \cdots, F_n$$

除去变数 $x_1,\cdots,x_n$ 外，如果还有另一些变数 $y_1,\cdots,y_l$，则$\frac{D(F_1,\cdots,F_n)}{D(x_1,\cdots,x_n)}\equiv 0$ 仅说明 $F_1,\cdots,F_n$ 间有函数关系，但这个关系可能还依数 $y_1,\cdots,y_l$.

附记 2　如果 X, Y 是 x, y, z 的函数，而 x, y, z 是 u, v 的函数，则

$$\frac{D(X,Y)}{D(u,v)}=\frac{D(X,Y)}{D(x,y)}\frac{D(x,y)}{D(u,v)}+\frac{D(X,Y)}{D(y,z)}\frac{D(y,z)}{D(u,v)}$$
$$+\frac{D(X,Y)}{D(z,x)}\frac{D(z,x)}{D(u,v)}.$$

应用：对数性质的直接证明：假定 $f(x)$ 是一个函数，当 $x=1$ 时为 0，而其微商等于 $\frac{1}{x}$. 令

$$u=f(x)+f(y),\quad v=xy,$$

则

$$\frac{D(u,v)}{D(x,y)}=\begin{vmatrix}\frac{1}{x} & \frac{1}{y}\\ y & x\end{vmatrix}=0.$$

因此有一个关系式

$$f(x)+f(y)=\varphi(xy).$$

取 $y=1$，则得 $f(x)=\varphi(x)$. 因此得对数性质.

§12. 代 数 处 理

在讲代数处理之前先来一个对比：

$$y=xA. \tag{1}$$

此处 $y=y^{(1,m)}$，$x=x^{(1,n)}$，$A=A^{(n,m)}$. (1) 代表 n 个变数 $x_1,\cdots,x_n$ 的 m 个线性方程组. 假定 A 的秩等于 r.

重新编排 y 及 x 的分量，可以使 A 中左上角的 r 行列的子行列式不等于 0，即

$$(y_1,\cdots,y_r,\ y_{r+1},\cdots,y_m)=(x_1,\cdots,\ x_r,\ x_{r+1},\cdots,x_n)\begin{pmatrix}A_1 & A_2\\ A_3 & A_4\end{pmatrix}. \tag{2}$$

此处 $A_1=A_1^{(r)},\ A_2=A_2^{(r,m-r)},\ A_3=A_3^{(n-r,r)},\ A_4=A_4^{(n-r,m-r)}$. 由此推得

$$(y_1,\cdots y_r;x_{r+1},\cdots,x_n)=(x_1,\cdots x_r,x_{r+1},\cdots,x_n)\begin{pmatrix}A_1 & 0\\ A_3 & I^{(n-r)}\end{pmatrix}, \tag{3}$$

代入上式得

$$(y_1,\cdots,y_r,y_{r+1},\cdots,y_m)=(y_1,\cdots,y_r,x_{r+1},\cdots,x_n)\begin{pmatrix}A_1 & 0\\ A_3 & I^{(n-r)}\end{pmatrix}^{-1}\begin{pmatrix}A_1 & A_2\\ A_3 & A_4\end{pmatrix}. \tag{4}$$

由于

$$\begin{pmatrix}A_1 & 0\\ A_3 & I^{(n-r)}\end{pmatrix}^{-1}=\begin{pmatrix}A_1^{-1} & 0\\ -A_3A_1^{-1} & I\end{pmatrix}$$

及 A 的秩等于 r 可知

$$\begin{pmatrix}A_1 & 0\\ A_3 & I\end{pmatrix}^{-1}\begin{pmatrix}A_1 & A_2\\ A_3 & A_4\end{pmatrix}=\begin{pmatrix}I^{(r)} & A_1^{-1}A_2\\ 0 & 0\end{pmatrix}.$$

因此

$$(y_1,\cdots,y_r,y_{r+1},\cdots,y_m)=(y_1,\cdots,y_r,x_{r+1},\cdots,x_n)\begin{pmatrix}I & A_1^{-1}A_2\\ 0 & 0\end{pmatrix}. \tag{5}$$

即

$$(y_{r+1},\cdots,y_m)=(y_1,\cdots,y_r)A_1^{-1}A_2. \tag{6}$$

即 $y_{r+1},\cdots,y_m$ 可表为 $y_1,\cdots,y_r$ 的线性函数.

现在来处理上节所提出的问题

$$du=dx\,\frac{\partial(u_1,\cdots,u_m)}{\partial(x_1,\cdots,x_n)}. \tag{1'}$$

重新编排 u 与 x 的次序，可以命 $\dfrac{\partial(u_1,\cdots,u_m)}{\partial(x_1,\cdots,x_n)}$ 中左上角的 r 行列的子行列式 $\neq 0$. 即

$$(du_1,\cdots,du_r,du_{r+1},\cdots,du_m)=(dx_1,\cdots,dx_r,dx_{r+1},\cdots,dx_n)\begin{pmatrix}A_1 & A_2\\ A_3 & A_4\end{pmatrix}, \tag{2'}$$

此处

$$A_1=\frac{\partial(u_1,\cdots,u_r)}{\partial(x_1,\cdots,x_r)},\ 等.$$

由此推出

$$(du_1,\cdots,du_r,dx_{r+1},\cdots,dx_n)=(dx_1,\cdots,dx_r,\ dx_{r+1},\cdots,dx_n)\begin{pmatrix}A_1 & 0\\ A_3 & I^{(n-r)}\end{pmatrix}. \tag{3'}$$

同时处理得

$$(du_1,\cdots,du_r,du_{r+1},\cdots,du_m)=(du_1,\cdots,du_r,\ dx_{r+1},\cdots,dx_n)\begin{pmatrix}I & A_1^{-1}A_2\\ 0 & 0\end{pmatrix}. \tag{4'}$$

由 Jacobian 乘法法则可知

$$\begin{pmatrix}I & A_1^{-1}A_2\\ 0 & 0\end{pmatrix}=\frac{\partial(u_1,\cdots,u_r,\ u_{r+1},\cdots,u_m)}{\partial(u_1,\cdots,u_r,\ x_{r+1},\cdots,x_n)} \tag{5}$$

即

$$\frac{\partial u_{r+1}}{\partial x_{r+1}}=\cdots=\frac{\partial u_{r+1}}{\partial x_n}=0. \tag{6}$$

即 u_{r+1} 与 $x_{r+1},\cdots,x_n$ 无关，同样 $u_{r+2},\cdots,u_m$ 都与 $x_{r+1},\cdots,x_m$ 无关．因此 $u_{r+1},\cdots,u_m$ 是 $u_1,\cdots,u_r$ 的函数．即

$$\begin{cases}u_{r+1}=v_{r+1}(u_1,\cdots,u_r),\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ u_m=v_m(u_1,\cdots,u_r)\ .\end{cases} \tag{7}$$

我们现在研究这 $m-r$ 个函数关系之外，是否还有其他．显然是有的，例如

$$u_{r+1}+u_{r+2}=v_{r+1}+v_{r+2},$$
$$u_{r+1}^3+7u_m^8=v_{r+1}^3+7v_m^8$$

等．因此，我们还必须明确所谓其他的函数关系．把(7)写成为

$$\begin{cases}r_1(u_1,\cdots,u_m)=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ r_p(u_1,\cdots,u_m)=0,\quad p=n-r.\end{cases} \tag{8}$$

由 (7) 可知这 p 个函数的函数行列式之秩等于 p.

如果 $u_1,\cdots,u_m$ 适合 (8)，则显然也适合

$$F(r_1,\cdots,r_p)=F(0,\cdots,0). \tag{9}$$

因此如果一个函数依赖于 $r_1,\cdots,r_p$，则也为 $u_1,\cdots,u_m$ 所适合．

所以我们的具体要求应当是：在(8)之外，是否还有

$$r_{p+1}(u_1,\cdots,u_m)=0$$

不与 $r_1,\cdots,r_p$ 函数相关的，也就是以(7)代入它不恒等于 0 的．回答是：没有了．如果有从(7)代入得到一个 $u_1,\cdots,u_r$ 的函数．即 $u_1,\cdots,u_r$ 函数相关，这和原先假定的

$$\frac{\partial(u_1,\cdots,u_r)}{\partial(x_1,\cdots,x_r)}\neq 0$$

相违背．

因此得出上节开始所谈到的结论．

请读者比较这两个处理方法：先看出来龙去脉，再看加工后的优点．

习题 1

$$u_1=\frac{x_1}{\sqrt{1-x_1^2-\cdots-x_n^2}},\cdots,\quad u_n=\frac{x_n}{\sqrt{1-x_1^2-\cdots-x_n^2}},$$

则

$$\frac{\partial(u_1,\cdots,u_n)}{\partial(x_1,\cdots,x_n)}=\frac{1}{(1-x_1^2-x_2^2\cdots-x_n^2)^{1+n/2}}.$$

习题 2

$$x_1=\cos\varphi_1,$$
$$x_2=\sin\varphi_1\cos\varphi_2,$$
$$x_3=\sin\varphi_1\sin\varphi_2\cos\varphi_3,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$x_n=\sin\varphi_1\sin\varphi_2\cdots\sin\varphi_{n-1}\cos\varphi_n,$$

则

$$\frac{D(x_1\cdots x_n)}{D(\varphi_1\cdots\varphi_n)}=(-1)^n\sin^n\varphi_1\sin^{n-1}\varphi_2\sin^{n-2}\varphi_3\cdots\sin^2\varphi_{n-1}\sin\varphi_n.$$

§ 13. Wronskian

n 个单变数的函数 $\phi_i(x)$，$i=1,2,\cdots,n$ 称为线性相关，如果有一组不全为 0 的常数 C_i 使

$$\sum_{i=1}^{n}C_i\phi_i(x)\equiv 0. \tag{1}$$

假定这些函数里可以微分 $(n-1)$ 次的，则微分 (1) 可得

$$\sum_{i=1}^{n}C_i\phi_i^{(j-1)}(x)\equiv 0,\quad j=1,2,\cdots,n.$$

消去 $C_1,\cdots,C_n$ 即得

$$W(\phi_1,\cdots,\phi_n)=\begin{vmatrix}\phi_1, & \cdots, & \phi_n,\\ \phi_1', & \cdots, & \phi_n'\\ \cdots & \cdots & \cdots\\ \phi_1^{(n-1)}, & \cdots, & \phi_n^{(n-1)}\end{vmatrix}\equiv 0.$$

$W(\phi_1,\cdots,\phi_n)$ 称为 Wronskian，这是线性相关的必要性，只在一些添加条件下，充分性才能正确.

我们假定 $W(\phi_1,\cdots,\phi_n)\equiv 0$. 并假定在某区间内 $W(\phi_1,\cdots,\phi_{n-1})\neq 0$. 由

$$\phi_n^{(j-1)}(x)=\sum_{i=1}^{n-1}u_i(x)\phi_i^{(j-1)}(x),\quad j=1,2,\cdots,n, \tag{2}$$

可以算出 $u_i(x)$ (唯一决定)，微分(2)式得

$$\phi_n^{(j)}(x)=\sum_{i=1}^{n-1}u_i(x)\phi_i^{(j)}(x)+\sum_{i=1}^{n-1}u_i'(x)\phi_i^{(j-1)}(x), \tag{3}$$

由 (2) 可知

$$0=\sum_{i=1}^{n-1}u_i'(x)\phi_i^{(j-1)}(x),\quad j=1,2,\cdots,n-1.$$

由于 $W(\phi_1,\cdots,\phi_{n-1})\neq 0$，所以 $u_i'(x)\equiv 0$. 即 $u_i(x)=C_i$. 因此

$$\phi_n(x)=\sum_{i=1}^{n-1}C_i\phi_i(x).$$

故得

定理 1 在 $a\leqslant x\leqslant b$ 中假定 $W(\phi_1,\cdots,\phi_n)\equiv 0$，并假定在其中所有的 $n-1$ 个函数的 Wronskian 不同时为 0，则 $\phi_1,\cdots,\phi_n$ 线性相关.

Peano 之反例:

$$\phi_1(x)\equiv 2x^2,\quad \phi_2(x)\equiv x^2,\qquad (x\geqslant 0);$$
$$\phi_1(x)\equiv x^2,\quad \phi_2(x)\equiv 2x^2,\qquad (x\leqslant 0).$$

则在 $x \geqslant 0$ 时，$\phi_1(x) = 2\phi_2(x)$，而 $x \leqslant 0$ 时，$\phi_1(x) = \frac{1}{2}\phi_2(x)$.

此定理在含有 $x = 0$ 的区间内不正确，其原因在于当 $x = 0$ 时，$W(\phi_1) = W(\phi_2) = 0$ 不合要求.

但如果加上解析条件则有

定理 2 命 $\phi_1(x), \cdots, \phi_n(x)$ 是个解析函数，它们是线性相关的必要且充分条件是 $W(\phi_1, \cdots, \phi_n) \equiv 0$.

第三章　方阵的函数、贯及级数

§1. 方阵的相似性

我们不再复习初等因子等，而仅将方阵相似性的定义与 Jordan 标准型简述如下：

定义　两个 n 行列的复元素方阵 A, B 称为相似，如果有一个满秩的方阵 P 使

$$A = PBP^{-1}.$$

用符号 $A \overset{E}{=} B$ 记之.

显然有 (i) $A \overset{E}{=} A$, (ii) 若 $A \overset{E}{=} B$ 则 $B \overset{E}{=} A$; (iii) 若 $A \overset{E}{=} B$, $B \overset{E}{=} C$, 则 $A \overset{E}{=} C$.

定理 1　如果 A 的初等因子是

$$(\lambda-\lambda_1)^{q_1},\ (\lambda-\lambda_2)^{q_2}, \cdots\cdots, (\lambda-\lambda_s)^{q_s},$$

则 A 相似于

$$\begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & J_s \end{pmatrix}.$$

这儿

$$J_i (= J_i^{(q_i)}) = \begin{pmatrix} \lambda_i, & 1, & 0, & \cdots \\ 0, & \lambda_i, & 1, & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0, & 0, & 0, & \cdots \lambda_i \end{pmatrix}.$$

这是一般书上所说的 Jordan 标准型(或法式).

为了应用方便，当 $\lambda_i \neq 0$ 时，我们有时取以下的方阵代替 J_i,

$$\lambda_i J^{(q_i)}, J = J^{(q_i)} = \begin{pmatrix} 1、 & 1、 & 0, & \cdots, & 0 \\ 0, & 1, & 1, & \cdots, & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

要证明这点十分容易，因为 $\lambda_i J^{(q_i)} \overset{E}{=} J_i$, 更具体些因为

$$J_i = [1, \lambda_i, \lambda_i^2, \cdots, \lambda_i^{(q_i-1)}]\lambda_i J[1, \lambda_i, \cdots, \lambda_i^{(q_i-1)}]^{-1}.$$

以上所说的相似关系是在复数范围内进行的. 在有些应用中我们不希望出现复数. 希望实的经实的变化而变为实的也就是：

定义　两个实元素方阵 A, B 称为实相似，如果有一个实满秩方阵 P 使

$$PAP^{-1} = B.$$

当然我们可以从头研究起. 但是我们既然学会了定理1，我们希望利用它来处理这新的问题.

定理 2　如果实 A 的初等因子

$(\lambda-\lambda_1)^{q_1}, \cdots, (\lambda-\lambda_s)^{q_s}$, $\lambda_1, \cdots, \lambda_s$ 是实数

$(\lambda-\rho_1 e^{i\theta_1})^{p_1},\cdots,\ (\lambda-\rho_r e^{i\theta_r})^{p_r}$, r 对共轭复数 $\rho_j e^{i\theta_j}$,

$(\lambda-\rho_1 e^{-i\theta_1})^{p_1},\cdots,\ (\lambda-\rho_r e^{-i\theta_r})^{p_r}$,

则 A 实相似于

$$T=\begin{pmatrix} \lambda_1 J^{(q_1)} & 0 & 0\cdots\cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 J^{(q_2)} & 0\cdots\cdots & 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & & & & \\ 0 & 0 & 0\cdots\cdots & \lambda_s J^{(q_s)} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots\cdots & 0 & \rho_1 K_1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & & & & \\ 0 & 0 & \cdots\cdots & 0 & 0 & 0 & \cdots & \rho_r K_r \end{pmatrix}.$$

(当 $\lambda_i=0$ 时仍以 $J_{\lambda i}$ 代 $\lambda_i J$) 此处

$$K_i=\begin{pmatrix} \cos\theta_i J^{(p_i)} & \sin\theta_i J^{(p_i)} \\ -\sin\theta_i J^{(p_i)} & \cos\theta_i J^{(p_i)} \end{pmatrix}.$$

这样的标准型是实的.

(读者自证，虚的初等因子共轭成对).

证明　由于

$$\rho\begin{pmatrix} I & -iI \\ I & iI \end{pmatrix}\begin{pmatrix} \cos\theta J & \sin\theta J \\ -\sin\theta J & \cos\theta J \end{pmatrix}\begin{pmatrix} I & -iI \\ I & iI \end{pmatrix}^{-1}$$

$$=\frac{\rho}{2}\begin{pmatrix} e^{i\theta}J & -ie^{i\theta}J \\ e^{-i\theta}J & ie^{-i\theta}J \end{pmatrix}\begin{pmatrix} I & I \\ iI & -iI \end{pmatrix}=\begin{pmatrix} \rho e^{i\theta}J & 0 \\ 0 & \rho e^{-i\theta}J \end{pmatrix},$$

所以由定理 1 可知 A 与 T 是复相似的，即有一复方阵 Q 使

$$QAQ^{-1}=T. \tag{1}$$

这儿 A 与 T 都是实的. 我们现在证明一定有一个实的 Q 适合于 (1), 命

$$Q=V+iW.\quad \det(V+iW)\neq 0.$$

则得

$$VA=TV,\qquad WA=TW. \tag{2}$$

考虑

$$f(x)=\det(V+xW),$$

这是一个 x 的实系数多项式. 由于 $f(i)\neq 0$, 所以 $f(x)$ 的系数不都等于 0, 因此有一实数 x_0 使 $f(x_0)\neq 0$, 而

$$P=V+x_0W$$

即合所求. 因为由 (2) 可知

$$(V+x_0W)A=T(V+x_0W),$$

即得

$$PAP^{-1}=T.$$

§2. 方 阵 的 幂

命

$$J = J^{(n)} = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots \\ 0 & 1 & 1 & 0 & \cdots \\ 0 & 0 & 1 & 1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix},$$

即 $a_{ii} = 1$，$a_{ii+1} = 1$，其他的 a_{ij} 都等于 0.

定理 1

$$J^l = \begin{pmatrix} 1 & l & \binom{l}{2} & \cdots & \binom{l}{n-1} \\ 0 & 1 & l & \cdots & \\ 0 & 0 & 1 & \cdots & \\ \cdots & \cdots & \cdots & \cdots & \end{pmatrix},$$

即 $a_{ii}^{(l)} = 1$，$a_{ii+1}^{(l)} = \binom{l}{1}$，$a_{i,i+2}^{(l)} = \binom{l}{2}$，……主对角线以下的元素$=0$.

证明　由

$$\binom{l}{i} + \binom{l}{i-1} = \binom{l+1}{i}$$

及归纳法假定 $J^{l+1} = J^l \cdot J$，

$$a_{ii+p}^{(l+1)} = \sum_{k=1}^{n} a_{ik}^{(l)} a_{ki+p} = a_{ii+p}^{(l)} + a_{i,i+p-1}^{(l)}$$

$$= \binom{l}{p} + \binom{l}{p-1} = \binom{l+1}{p},$$

故得所云.

定理 2

$$\begin{pmatrix} \cos\theta J & \sin\theta J \\ -\sin\theta J & \cos\theta J \end{pmatrix}^l = \begin{pmatrix} \cos l\theta J^l, & \sin l\theta J^l \\ -\sin l\theta J^l, & \cos l\theta J^l \end{pmatrix}.$$

§ 3. 方阵乘幂的极限

先看方阵乘幂的极限

$$\lim_{l\to\infty} M^l \tag{1}$$

存在的条件是什么？一个矩阵其 M_l 当 $l \to \infty$ 时趋于极限 M_0 的意义是 M_l 的第 i 行 j 列的元素 $m_{ij}^{(l)}$ 当 $l \to \infty$ 时趋于 M_0 对应的元素 $m_{ij}^{(0)}$.

看 (1) 的极限与看与M相似的N是否趋于极限完全相同. 因此先看 Jordan 方阵趋限情况.

1) 特征值 λ 的绝对值$|\lambda| > 1$ 的情况，由于 $\lim\limits_{\lambda\to\infty}\lambda^l$ 不存在，所以如M有一个特征根它的绝对值>1，则 (1) 不存在.

2) 如果 $|\lambda| = 1$ 而 $\lambda \neq 1$，由于

$$\lim_{l\to\infty} e^{il\alpha},\quad 0 < \alpha < 2\pi$$

不存在，因而如果M有一个非 1 而绝对值等于 1 的特征根，则 (1) 仍然不存在.

3）如果 1 是特征根，但并非单的初等因子，由于

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^l = \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & \infty \\ 0 & 1 \end{pmatrix}$$

也不存在.

总之，如果要求 (1) 存在M只能有绝对值小于 1 的特征根及等于 1 而是单的初等因子的特征根.

4）再看 $|\lambda| < 1$ 的情况.

$$(\lambda J)^l = \lambda^l \begin{pmatrix} 1, \binom{l}{1}, \cdots \\ 0, \quad 1 \quad \cdots \\ \qquad\qquad \ddots \end{pmatrix},$$

由于 $\lambda^l \to 0$，$l\lambda^l \to 0$，$\binom{l}{2}\lambda^l \to 0$，……. 所以 $(\lambda J)^l \to 0$ （当 $\lambda = 0$ 时，$J_\lambda^n = 0$，不必证明).

因此得

定理 1 如果 A 仅有绝对值不大于 1 的特征根，而绝对值等于 1 的特征值只有 1，并且它的初等因子是单的. 这样

$$\lim_{l\to\infty} A^l = A_0$$

是存在的. 并且只有这样的情况是存在的.

这极限 A_0 的秩等于 A 中特征根 1 的重数. 而且 $A_0^2 = A_0$ (是一幂等元素).

定理 2 在定理 1 的假定下如果 1 是 A 的单根，则

$$\lim_{l\to\infty} A^l = u'v, \quad vu' = 1.$$

而且 u' 与 u 是 A 的特征值 1 的特征矢量(列或行). 并且除一常数因子外，这是唯一的.

证明　由定理 1 可知 A_0 的秩等于 1，因此可以写成为 $u'v$，这儿 u，v 是二矢量. 又由于幂等性质，即

$$u'v = u'vu'v = (vu')(u'v).$$

即

$$(1 - vu')u'v = 0.$$

而 $u'v \neq 0$，所以 $vu' = 1$.

再证 v 是对应于 A 的特征根 1 的特征 (行) 矢量. 假定 v_1 是对应于 A 的特征根 1 的特征矢量，即

$$v_1 A = v_1,$$

连用 l 次，得

$$v_1 A^l = v_1.$$

当 $l \to \infty$，则得

$$v_1 u' v = v_1.$$

因此

$$(v_1 u')v = v_1.$$

即 v 与 v_1 相差一个常数因子 $v_1 u'$，而且 $vA = v$.

同法证明

$$Au' = u'.$$

定理 3（极限定理） 仍如定理 2 的假定，对任一矢量 x，作矢量

$$xA^l = x_l,$$

则 x_l 的极限仅与 v 相差一个常数因子. 也就是不管原始假定如何最终结果的矢量各支量的比例恒不变.

证明 由

$$\lim_{l\to\infty} xA^l = xu'v = (xu')v$$

而得.

§4. 幂 级 数

现在研究幂级数

$$\sum_{l=0}^{\infty} a_l X^l, \tag{1}$$

这儿 a_l 是复数，X 是一变 n 行列方阵.

假定

$$\sum_{l=0}^{\infty} a_l x^l, \tag{2}$$

的收敛半径是 ρ 即当 $|x| < \rho$ 时(2)收敛.

这问题的研究也就跟 Jordan 标准型代 X 而得的问题等价.

1) 已知当 $|x| < \rho$ 时，

$$\sum a_l x^l,\ \sum a_l \binom{l}{1} x^l,\ \sum a_l \binom{l}{2} x^l, \cdots,\ \sum a_l \binom{l}{n} x^l, \cdots$$

都收敛. 所以如果 X 的所有的特征根的绝对值都 $< \rho$ 时，(1) 也收敛.

2) 当 X 有一个特征根的绝对值 $> \rho$，则 (1) 发散.

3) 当 X 的所有的特征根的绝对值都 $\leqslant \rho$，关于绝对值 $= \rho$ 的特征根 $\lambda_0 = \rho e^{i\theta}$，如果对应于 λ_0 的初等因子 $(\lambda - \lambda_0)^{l_1}, \cdots, (\lambda - \lambda_0)^{l_s}$，$l_1 \geqslant l_2 \geqslant \cdots \geqslant l_s$，则需要看

$$f(\lambda_0) = \sum_{l=0}^{\infty} a_l (\rho e^{i\theta})^l,$$

$$f'(\lambda_0) = \sum_{l=1}^{\infty} l a_l (\rho e^{i\theta})^{l-1},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$f^{(l_1)}(\lambda_0) = \sum_{l=l_1}^{\infty} l(l-1)\cdots(l-l_1+1) a_l (\rho e^{i\theta})^{l-l_1}$$

是否收敛而定. 如果都收敛，则(1)收敛，有一发散，则 (1) 发散.

§5. 幂级数举例

1)

$$\exp X = e^X = \sum_{l=0}^{\infty} \frac{1}{l!} X^l$$

这个级数对任意的方阵X都收敛.

如果 $XY=YX$，则

$$e^X\cdot e^Y=\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\frac{X^lY^m}{l!m!}=\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{l=0}^{n}\frac{n!}{l!(n-l)!}X^lY^{n-l}$$
$$=\sum_{n=0}^{\infty}\frac{(X+Y)^n}{n!}=e^{X+Y}.$$

但当 $XY\neq YX$ 时，此式不能成立.

假定X的特征根等于 $\lambda_1,\cdots,\lambda_n$，则 e^X 的特征根等于

$$e^{\lambda_1},\cdots\cdots,e^{\lambda_n}.$$

由于(当 $\lambda\neq 0$ 时，取 $J=J^{(s)}$)

$$\exp\lambda J=\sum_{l=0}^{\infty}\frac{\lambda^l}{l!}J^l=\sum_{l=0}^{\infty}\frac{\lambda^l}{l!}\begin{bmatrix}1 & l & \frac{l(l-1)}{2} & \cdots & \\ 0 & 1 & l & \ddots & \\ & \ddots & \ddots & \ddots & l\\ & & & \ddots & 1\end{bmatrix}=\begin{pmatrix}e^\lambda & \lambda e^\lambda & \frac{\lambda^2}{2}e^\lambda\cdots\\ & & e^\lambda\end{pmatrix},$$

所以对应的初等因子等于 $(x-e^\lambda)^s$.

当 $\lambda=0$ 时 $\exp J_0$ 实际上是一个多项式，而初等因子是 $(x-1)^s$.

2）同样可以定义

$$\cos X=\frac{1}{2}(e^{iX}+e^{-iX})=\sum_{l=0}^{\infty}(-1)^l\frac{X^{2l}}{(2l)!},$$
$$\sin X=\frac{1}{2i}(e^{iX}-e^{-iX})=\sum_{l=0}^{\infty}(-1)^l\frac{X^{2l+1}}{(2l+1)!}.$$

不难直接证明

$$\cos^2X+\sin^2X=I.$$

同样可定义

$$\cosh X=\frac{1}{2}(e^X+e^{-X})=\sum_{l=0}^{\infty}\frac{X^{2l}}{(2l)!},$$
$$\sinh X=\frac{1}{2}(e^X-e^{-X})=\sum_{l=0}^{\infty}\frac{X^{2l+1}}{(2l+1)!}.$$

3）对数函数

$$\log(I+X)=X-\frac{X^2}{2}+\frac{X^3}{3}-\cdots\cdots$$

这个级数当X的所有的特征根都小于1时收敛. 不难证明，在同样条件下:

$$e^{\log(I+X)}=I+X.$$

证法，先由 $e^{\log(1+x)}=1+x$ 开始得出一些恒等式然后再证上式.

4）二项式展开

$$(I+X)^\alpha=I+\alpha X+\frac{\alpha(\alpha-1)}{2}X^2+\frac{\alpha(\alpha-1)(\alpha-2)}{3!}X^3+\cdots$$

也是当X的所有特征根都小于 1 时收敛.

§6. 迭 代 法

看上法

$$(I-A)^{-1}=I+A+A^2+\cdots \tag{1}$$

是一个极简单的公式，但是它是计算数学上迭代法的根源. 当A的特征根的绝对值都小于 1 时. 级数 (1) 收敛.

迭代法是：求联立方程组

$$x(I-A)=b \tag{2}$$

的解有以下的方法. 取 $x_0=b$，逐步地求出

$$x_i=b+x_{i-1}A,\quad i=1,2,\cdots\cdots.$$

这样递代的结果 $\lim\limits_{i\to\infty}x_i$ 是方程组 (1) 的解.

证明此点十分容易. 由归纳法不难证明

$$x_i=b(I+A+\cdots+A^i).$$

因此

$$\lim_{i\to\infty}x_i=b(I-A)^{-1}.$$

怎样判断，A的特征根的绝对值是否小于 1？以下的条件可供参考. 如果

$$\sum_{i=1}^{n}|a_{ij}|\leqslant g<1, \tag{3}$$

则级数 (1) 收敛.

证明 命 $A^l=(a_{ij}^{(l)})$，今往证 $\sum\limits_{i=1}^{n}|a_{ij}^{(l)}|\leqslant q^l$，用归纳法

$$\sum_{i=1}^{n}|a_{ij}^{(l+1)}|=\sum_{i=1}^{n}\left|\sum_{k=1}^{n}a_{ik}^{(l)}a_{kj}\right|\leqslant\sum_{k=1}^{n}\left(\sum_{i=1}^{n}|a_{ik}^{(l)}|\right)|a_{kj}|$$

$$\leqslant q^l\sum_{k=1}^{n}|a_{kj}|\leqslant q^{l+1}.$$

所以

$$\sum_{l=0}^{\infty}|a_{ij}^{(l)}| \text{ 收敛.}$$

附记 1 这定理不但证明了收敛，而且给出了误差. 误差是

$$\sum_{l=n+1}^{\infty}|a_{ij}^{(l)}|\leqslant\sum_{l=n+1}^{\infty}q^l=\frac{q^{n+1}}{1-q}.$$

附记 2 一般的特征值绝对值都小于 1 的方阵 A，可以经过以下的处理使其适合 (3). 可以找到一个对角线方阵 $\Lambda=[\lambda_1,\cdots,\lambda_n]$ 使

$$\Lambda A\Lambda^{-1}$$

适合于条件 (3). 这样 $\lambda_1,\cdots,\lambda_n$ 的存在性可以证之如次. 命 ρ 是A的绝对值最大的特征根之一. 而 $(\lambda_1,\cdots,\lambda_n)$ 是它所对应的特征矢量

$$\sum_{i=1}^{n}\lambda_i a_{ij}=\rho\lambda_j.$$

写成方阵形式

$$[\lambda_1,\cdots,\lambda_n](a_{ij})[\lambda_1,\cdots,\lambda_n]^{-1}=(b_{ij}),$$

而

$$\sum_{i=1}^{n}b_{ij}=\rho.$$

在一般计算时,不一定要真的找出 $\lambda_1,\cdots,\lambda_n$ 来．而只要找出 $\lambda_1,\cdots,\lambda_n$ 使 $\sum_{i=1}^{n}b_{ij}<1$ 即可.

附记 3　这个迭代法当然对 $I-A$ 求逆也行．即命

$$B_0=I,\quad B_l=I+AB_{l-1}.$$

这个迭代过程还可以考虑以下的收敛得更快的形式

A	$T_1=I+A$
A^2	$T_2=T_1+A^2T_1$
$A^4=(A^2)^2$	$T_3=T_2+A^4T_2$
$A^8=(A^4)^2$	$T_4=T_3+A^8T_3$
······	·········

§7. 关于指数函数

定理 1
$$\lim_{n\to\infty}\left(I+\frac{1}{n}A\right)^n=e^A.$$

证明　由于

$$n\log\left(I+\frac{1}{n}A\right)=A-\frac{1}{2n}A^2+\frac{1}{3n^2}A^3+\cdots,$$

因此

$$\lim_{n\to\infty}n\log\left(I+\frac{1}{n}A\right)=A.$$

即得定理

定理 2　一个满秩方阵一定是一个指数函数的值.

证明　1）先考虑最特殊的情况:

$$J=\begin{pmatrix}1 & 1 & 0\cdots 0\\ & \cdots & \\ 0 & & 1\end{pmatrix}=I+L\ (J=J^{(q)}).$$

由于 $L^q=0$，因此

$$\log J=\log(I+L)=L-\frac{1}{2}L^2+\frac{1}{3}L^3\cdots,$$

因此

$$\log J = \begin{pmatrix} 0 & 1 & -\frac{1}{2} & \frac{1}{3} \cdots\cdots \\ & & & \\ 0 & & & \end{pmatrix} = K.$$

即 $J = e^K$.

2）由上推得

$$\lambda J = e^{\log \lambda I + K}.$$

3）不难推到一般的满秩方阵.

定理 3 e^{At} 的所有的元素非负的必要且充分条件是

$$a_{ij} \geqslant 0, \qquad i \neq j.$$

证明 1）先考虑 $a_{ij} > 0$ 的情况由于

$$e^{At} = I + At + \frac{1}{2} A^2 t^2 + \cdots, \quad (t > 0),$$

所以当 t 充分小时 e^{At} 的所有的元素都是正的（对角线上的元素与 I 的对应元素同号，非对角线上的元素与 At 的对应元素同号）. 由于

$$e^{At} = (e^{At/n})^n$$

可知 e^{At} 的元素也是正的.

2）其次由连续性处理 $a_{ij} \geqslant 0$ 的情况.

3）反之，如果 $a_{12} < 0$，则在

$$e^{At} = I + At + \frac{1}{2!} A^2 t^2 + \cdots = (b_{ij})$$

中，当 t 充分小时，$b_{12} < 0$，因而得出本定理.

§8. 单变数方阵的微分运算

假定 $A(t) = (a_{ij}(t))$ 是 n 行列的方阵，t 是变数. 如果存在，则我们定义

$$\frac{d}{dt} A(t) = \lim_{h \to 0} \frac{1}{h} (A(t+h) - A(t)) = \left(\frac{d}{dt} a_{ij}(t) \right)$$

微分的一些性质如下：

(i)
$$\frac{d}{dt}(A + B) = \frac{d}{dt} A + \frac{d}{dt} B.$$

(ii)
$$\frac{d}{dt}(AB) = A \frac{dB}{dt} + \frac{dA}{dt} B.$$

如果 A 是满秩的，立刻推得（取 $B = A^{-1}$）

(iii)
$$\frac{d}{dt}(A^{-1}) = -A^{-1} \frac{dA}{dt} A^{-1}.$$

但切需注意 $\frac{dA}{dt}$ 与 A 不一定可交换，因此

$$\frac{dA^2}{dt} = A \frac{dA}{dt} + \frac{dA}{dt} A, \qquad \frac{dA^3}{dt} = A^2 \frac{dA}{dt} + A \frac{dA}{dt} A + \frac{dA}{dt} A^2, \text{ 等.}$$

幂级数

$$F(t)=A_0+A_1t+A_2t^2+A_3t^3+\cdots,$$

在收敛范围内可以逐项求微分得

$$\frac{dF(t)}{dt}=A_1+2A_2t+3A_3t^2+\cdots.$$

例如:

(iv) $$\frac{d}{dt}e^{At}=Ae^{At}.$$

(v) $$\frac{d}{dt}\log(I+At)=\frac{d}{dt}\left(At-\frac{1}{2}A^2t^2+\frac{1}{3}A^3t^3-\cdots\right)$$
$$=A(I-At+A^2t^2-\cdots)=A(I+At)^{-1}$$

等.

积分也定义为

$$\int_c A(t)\,dt=\int_c(a_{ij}(t))\,dt=\left(\int_c a_{ij}(t)\,dt\right).$$

这积分可以作为实变数在一区间上的,也可以作为复变数在复平面的一个曲线上的.

由 (ii) 推得分部积分公式

$$\int_c A\frac{dB}{dt}dt+\int_c\frac{dA}{dt}B\,dt=AB\Big|_c.$$

第三章的补充

§1. Jordan 标准型的幂级数

1）命

$$J_\lambda=\begin{pmatrix}\lambda & 1 & 0 & \cdots & 0\\ 0 & \lambda & 1 & \cdots & 0\\ 0 & 0 & \lambda & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & \lambda\end{pmatrix},$$

则

$$J_\lambda^p=\begin{pmatrix}\lambda^p, & p\lambda^{p-1}, & \frac{p(p-1)}{2!}\lambda^{p-2}, & \cdots\\ 0, & \lambda^p, & p\lambda^{p-1}, & \cdots\\ \cdots & \cdots & \cdots & \cdots\\ 0, & 0, & 0, & \cdots\cdots\end{pmatrix}$$

$$=\begin{pmatrix}\lambda^p, & (\lambda^p)', & \frac{1}{2!}(\lambda^p)'', & \cdots\\ 0, & \lambda^p, & (\lambda^p)', & \cdots\\ \cdots & \cdots & \cdots & \cdots\end{pmatrix}.$$

证明　用归纳法，由于$(\lambda\cdot\lambda^p)^{(t)}=\lambda(\lambda^p)^{(t)}+t(\lambda^p)^{(t-1)}$，所以

$$\frac{1}{t!}\lambda(\lambda^p)^{(t)}+\frac{1}{(t-1)!}(\lambda^p)^{(t-1)}=\frac{1}{t!}(\lambda^{p+1})^{(t)}.$$

因此得到以上的结果.

2）命

$$f(x)=\sum_{p=0}^{\infty}a_px^p,$$

则

$$f(J_\lambda)=\sum_{p=0}^{\infty}a_p\begin{pmatrix}\lambda^p, & (\lambda^p)', & \frac{1}{2}(\lambda^p)'', & \cdots\\ 0, & \lambda^p, & (\lambda^p)', & \cdots\\ \cdots & \cdots & \cdots & \cdots\end{pmatrix}$$

$$=\begin{pmatrix}f(\lambda), & f'(\lambda), & \frac{1}{2}f''(\lambda), & \cdots\\ 0, & f(\lambda), & f'(\lambda), & \cdots\\ 0, & 0, & f(\lambda), & \cdots\\ \cdots & \cdots & \cdots & \cdots\end{pmatrix}.$$

3）如果

$$X = P(J_{\lambda_1} \dotplus J_{\lambda_2} \dotplus \cdots \dotplus J_{\lambda_s})P^{-1},$$

则

$$f(X) = P(K_{\lambda_1} \dotplus K_{\lambda_2} \dotplus \cdots \dotplus K_{\lambda_s})P^{-1}$$

这儿

$$K_{\lambda i} = \begin{pmatrix} f(\lambda_i) & f'(\lambda_i), & \frac{1}{2}f''(\lambda_i), & \cdots \\ 0 & f(\lambda_i), & f'(\lambda_i), & \cdots \\ 0, & 0, & f(\lambda_i), & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

4）特例有

$$e^{J}\lambda = \begin{pmatrix} e^{\lambda}, & e^{\lambda}, & \frac{1}{2}e^{\lambda} & \cdots \\ 0, & e^{\lambda}, & e^{\lambda} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix},$$

而

$$e^{tJ}\lambda = \begin{pmatrix} e^{t\lambda}, & te^{t\lambda}, & \frac{1}{2}t^2e^{t\lambda}, & \cdots \\ 0, & e^{t\lambda}, & te^{t\lambda}, & \cdots \\ 0, & 0, & e^{t\lambda}, & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

§2. 数的方阵幂

1）我们定义

$$x^A = e^{A\log x} = \sum_{n=0}^{\infty} \frac{(A\log x)^n}{n!},$$

这时

$$x^A \cdot x^B$$

不一定等于

$$x^{A+B}.$$

微商规则

$$\frac{d}{dx}x^A = \frac{d}{dx}e^{A\log x} = \frac{d}{dx}\sum_{l=0}^{\infty}\frac{(A\log x)^l}{l!}$$

$$= \frac{1}{x}\sum_{l=0}^{\infty}\frac{(A\log x)^{l-1}A}{(l-1)!} = Ax^{A-I}(=x^{A-I}A),$$

这儿用了

$$x^{-I} = e^{-I\log x} = \sum_{n=0}^{\infty}\frac{(-I\log x)^n}{n!} = \sum_{n=0}^{\infty}\frac{(-\log x)^n}{n!}I$$

$$= x^{-1}I.$$

2）
$$x^{J\lambda}=e^{J\lambda\log x}=\begin{pmatrix} x^{\lambda}, & (\log x)x^{\lambda}, & \frac{(\log x)^2}{2!}x^{\lambda}, & \cdots \\ 0, & x^{\lambda}, & (\log x)x^{\lambda}, & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix},$$

如果 λ 的实数部份 > 0，则

$$\lim_{x\to 0} x^{J\lambda}=0.$$

3）**定理 1**　如果 A 的所有的特征根的实数部份都 > 0，则

$$\lim_{x\to 0} x^{A}=0.$$

证明　如果

$$A=PBP^{-1},$$

则

$$x^{A}=Px^{B}P^{-1}.$$

因此把问题归结为 A 为 Jordan 标准型的情况来讨论，由 2）知道本定理对 Jordan 标准型对，因而本定理也对.

§3. 特殊 X 的 e^X

1）先看 $n=2$，$X=-X'$，即

$$X=\begin{pmatrix} O & \lambda \\ -\lambda & O \end{pmatrix}.$$

如此则

$$X^2=-\lambda^2 I.$$

因此

$$e^X=\sum_{l=0}^{\infty}\frac{(-\lambda^2)^l}{(2l)!}I+\sum_{l=0}^{\infty}\frac{(-\lambda^2)^l\lambda}{(2l+1)!}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}=\begin{pmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{pmatrix}.$$

这便是平面上的旋转.

2）$n=3$.

三维空间的旋转是

$$\begin{pmatrix} \cos\alpha\cos\beta\cos\gamma-\sin\alpha\sin\gamma, & \cos\alpha\cos\beta\sin\gamma+\sin\alpha\cos\gamma, & -\cos\alpha\sin\beta \\ -\sin\alpha\cos\beta\cos\gamma-\cos\alpha\sin\gamma, & -\sin\alpha\cos\beta\sin\gamma+\sin\alpha\cos\gamma, & \sin\alpha\sin\beta \\ \sin\beta\cos\gamma, & \sin\beta\sin\gamma, & \cos\beta \end{pmatrix}$$

$$=\begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}\begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

因此它等于

$$\exp\begin{pmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\exp\begin{pmatrix} 0 & 0 & -\beta \\ 0 & 0 & 0 \\ \beta & 0 & 0 \end{pmatrix}\exp\begin{pmatrix} 0 & \gamma & 0 \\ -\gamma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

反过来，从一般的三行三列的斜对称方阵

$$X=\theta\begin{pmatrix}0 & -n & m\\ n & 0 & -l\\ -m & l & 0\end{pmatrix},\quad l^2+m^2+n^2=1$$

出发，由此得

$$X^2=-\theta^2(I-u'u).$$

此处 $u=(l, m, n)$, $uu'=l^2+m^2+n^2=1$. 由此得

$$X^{2l}=(-\theta^2)^l(I-u'u)^l=(-\theta^2)^l(I-u'u),\quad l=1,2,\cdots.$$

及

$$uX=0.$$

因此

$$e^X=I+\sum_{l=1}^{\infty}\frac{(-\theta^2)^l}{(2l)!}(I-u'u)+\sum_{l=0}^{\infty}\frac{(-\theta^2)^l}{(2l+1)!}(I-u'u)X$$

$$=\cos\theta I+(1-\cos\theta)u'u+\sin\theta X$$

$$=\begin{pmatrix}\cos\theta+l^2(1-\cos\theta), & -n\sin\theta+ml(1-\cos\theta), & m\sin\theta+nl(1-\cos\theta)\\ n\sin\theta+ml(1-\cos\theta), & \cos\theta+m^2(1-\cos\theta), & -l\sin\theta+mn(1-\cos\theta)\\ -m\sin\theta+nl(1-\cos\theta), & l\sin\theta+mn(1-\cos\theta), & \cos\theta+n^2(1-\cos\theta)\end{pmatrix}$$

这就是绕以 (l, m, n) 为方向余弦的轴转 θ 度的旋转公式.

3）更一般些，$X=-X'$，则

$$e^X\cdot(e^X)'=e^X\cdot e^{X'}=e^X\cdot e^{-X}=I.$$

即方阵

$$\Gamma=e^X$$

适合于

$$\Gamma\Gamma'=I.$$

即斜对称方阵的指数函数是正交方阵，而 X 称为正交方阵的"无穷小"方阵，如果 Γ 的无穷小方阵是 X，则 Γ^l 的无穷小方阵是 lX 也是

$$\lim_{t\to 0}\frac{I-\Gamma^t}{t}=-X.$$

反过来从 $\Gamma\Gamma'=I$ 出发，命 $\Gamma=e^X$，则由

$$e^X=\Gamma=(\Gamma')^{-1}=e^{-X'},$$

可推得 $X=-X'$.

4）假定 $trX=0$，即 X 的诸特征根之和等于 0，因此 e^X 诸特征根之积等于 1，即得：如果 $trX=0$，则

$$|e^X|=1.$$

5）命 $n=2l$

$$F=\begin{pmatrix}O^{(l)} & 1\\ -I & O\end{pmatrix}.$$

假定 X 适合于

$$FX+X'F=O,$$

则 $FX^n=(-1)^nX'^nF$，所以

$$Fe^X=\sum_{n=0}^{\infty}F\frac{X^n}{n!}=\sum_{n=1}^{\infty}\frac{(-X')^n}{n!}=e^{-X'}F.$$

命 $P = e^X$，则

$$P'FP = F.$$

6）如果

$$X = -\bar{X}',$$

则

$$e^X(e^{\bar{X}})' = e^X e^{-X} = I.$$

§ 4. e^X 与 X 的对应关系

1）给了 X，我们可以由

$$M(X) = e^X = \sum_{l=0}^{\infty} \frac{X^l}{l!}$$

得出 e^X. 另一方面，如果M非奇异，则可以表为 e^X，而且有关系

$$\lim_{t\to 0}\frac{M^t - I}{t} = \lim_{t\to 0}\frac{e^{Xt} - I}{t} = X.$$

由此显然可得

i）M^{-1} 对应于$-X$.

ii）M' 对应于 X'.

iii）$\bar{M}$ 对应于 $\bar{X}$.

M 的正交性，即 $M' = M^{-1}$，变为X的斜对称性 $X' = -X$，M的酉方阵性，即 $\bar{M}' = M^{-1}$，变为 iX 的 Hermite 性.

2）最重要的一个对应关系是

$$e^{sX}\cdot e^{tY}\cdot(e^{sX})^{-1}(e^{tY})^{-1},$$

所对应的元素是

$$\begin{aligned}
&\lim_{t\to 0}\lim_{s\to 0}\frac{e^{sX}\cdot e^{tY}\cdot e^{-sX}\cdot e^{-tX} - I}{st}\\
&= \lim_{t\to 0}\lim_{s\to 0}\frac{e^{sX} - I + e^{sX}e^{tY}(e^{-sX} - I)e^{-tY}}{st}\\
&= \lim_{t\to 0}\frac{X - e^{tY}Xe^{-tY}}{t}\\
&= \lim_{t\to 0}\frac{X(I - e^{-tY}) + (I - e^{tY})Xe^{-tY}}{t}\\
&= XY - YX.
\end{aligned}$$

第四章　常系数差分方程与常微分方程

§1. 差分方程

一个变数的差分方程通常是指：求出函数 $y=y(x)$ 适合于方程

$$F(y(x+n\delta),y(x+(n-1)\delta),\cdots,y(x+\delta),\ y(x),x)=0. \tag{1}$$

这儿 δ 是一个固定的数.

并不失去普遍性，在研究差分方程(1)时，我们假定 $\delta=1$. 并且假定我们已经解出

$$y(x+n)=\varphi(y(x+n-1),\cdots,y(x+1),y(x),x). \tag{2}$$

如果已经知道了

$$y(0)=c_0,y(1)=c_1,\cdots,y(n-1)=c_{n-1}. \tag{3}$$

则公式(2)便是一个递推公式，可以逐步地算出

$$y(n),\ y(n+1),\ y(n+2),\cdots$$

等值.

这样的差分方程称为 n 级的差分方程，所给的数值(3)称为初始值. 注意，当初始值定了，我们可以得出 y 在所有的自然数 n 时的数值，而并不能决定其他值，但如果给与了

$$y(\tau),y(\tau+1),\cdots,y(\tau+n-1), \tag{4}$$

则我们可以得出

$$y(\tau+l),\quad l=0,1,2,3,\cdots$$

诸值. 只有在 $0\leqslant\tau<1$ 中给了 n 个函数

$$y(\tau+p)=\varphi_p(\tau),\quad p=0,1,\cdots,n-1, \tag{5}$$

才能在整个直线上决定 $y(x)$，$x\geqslant 0$.

如果(2)还可以解出 $y(x)$ 的数值，则因而也可决定 $x\leqslant 0$ 时的 $y(x)$ 值.

差分方程的用场随着电子计算机而日益显著. 首先是在求微分方程的数值解法时，我们常用差分法，变为差分方程求解；其次，我们可以直接处理离散类型的问题，而不必经过微分方程，很可能原始数据是离散的，最后要求的解答也是离散性的，这样我们可以经过代数运算而直接处理，这是为什么本书中首先着重在线性方程的数值计算，其次差分方程，再其次是微分方程的道理.

把常微分方程

$$y^{(n)}(x)=\varphi(y^{(n-1)}(x),\cdots y'(x),y(x),x) \tag{6}$$

中的微商 $y^{(i)}(x)$ 代以未取极限之前的形式；即以

$$\frac{y(x+\delta)-y(x)}{\delta}\text{代 }y'(x)$$

$$\frac{y(x+2\delta)-2y(x+\delta)+y(x)}{\delta^2}\text{代 }y''(x)\text{ 等等.}$$

于是我们就得出一个差分方程.

这差分方程的解与 δ 有关，当 $\delta\to 0$ 时，一般它可能是微分方程(6)的解，我们先举几个例子来加以说明.

例 1　解一级齐次差分方程

$$y(x+1)=ay(x),\ y(0)=c.$$

由归纳法易证

$$y(n)=a^n c,\ n=0,1,2,\cdots. \tag{7}$$

再考虑一阶齐次微分方程

$$y'(x)=\lambda y(x),\ y(0)=k, \tag{8}$$

其差分形式是

$$y(x+\delta)-y(x)=\delta\lambda y(x),$$

即

$$y(x+\delta)=(1+\delta\lambda)y(x).$$

由差分方程的结果易见

$$y(n\delta)=k(1+\lambda\delta)^n.$$

命 $n\delta=x$，则

$$y(x)=k\left(1+\frac{x\lambda}{n}\right)^n.$$

当 $n\to\infty$ 时，

$$y(x)=ke^{\lambda x},$$

这就是微分方程的解.

例 2　解

$$y(x+1)=ay(x)+f(x),\quad y(0)=c.$$

命

$$y(x)=a^x c(x),\quad c(0)=c,$$

则得

$$c(x+1)=c(x)+a^{-x-1}f(x).$$

因此得出

$$c(l)=c+\sum_{p=1}^{l}a^{-p}f(p-1).$$

所以

$$y(l)=ca^l+\sum_{p=1}^{l}a^{l-p}f(p-1).$$

再考虑微分方程

$$\frac{dy}{dx}=\lambda y(x)+g(x),y(0)=c.$$

由此导出差分方程

$$y(x+\delta)=(1+\delta\lambda)y(x)+\delta g(x).$$

因此得

$$y(l\delta)=c(1+\delta\lambda)^l+\delta\sum_{p=1}^{l}(1+\delta\lambda)^{l-p}g((p-1)\delta).$$

命 $l\delta = x$,

$$y(x) = c\left(1 + \frac{x\lambda}{l}\right)^l + \frac{x}{l}\sum_{p=1}^{l}\left(1 + \frac{x\lambda}{l}\right)^{l-p} g((p-1)x/l)$$

$$= C\left(1 + \frac{x\lambda}{l}\right)^l + \left(1 + \frac{x\lambda}{l}\right)^l \cdot \frac{x}{l}\sum_{p=1}^{l} g\left(\frac{(p-1)x}{l}\right)e^{-p l g\left(1+\frac{x\lambda}{l}\right)},$$

当 $l \to \infty$ 时，$y(x)$ 趋于

$$e^{\lambda x}\left(C + \int_0^x g(t)e^{-\lambda t}dt\right).$$

这就是所讨论的微分方程的解.

附记，固然我们可以用

$$\frac{y(x+\delta) - y(x)}{\delta}$$

来代 $y'(x)$，但有时我们也可以用

$$\frac{y(x+\delta) - y(x-\delta)}{2\delta}$$

来代 $y'(x)$，一般讲来后者比前者还好些，原因是

$$\frac{y(x+\delta) - y(x-\delta)}{2\delta} - y'(x) \sim \frac{\delta^2}{6}y'''(x+\theta\delta)$$

$$\left(y(x+\delta) = y(x) + \delta y'(x) + \frac{\delta^2}{2}y''(x) + \frac{\delta^3}{6}y'''(x+\theta\delta)\right) \text{ 比}$$

$$\frac{y(x+\delta) - y(x)}{\delta} - y'(x) \sim \frac{\delta}{2}y''(x+\theta'\delta)$$

更精确些.

§2. 常系数线性差分方程——母函数法

解 n 级差分方程

$$a_0y(x+n) + a_1y(x+n-1) + \cdots + a_ny(x) = g(x), \tag{1}$$

其初始条件是

$$y(0) = c_0, \cdots, y(n-1) = c_{n-1}. \tag{2}$$

1) 先研究 $g(x) = 0$ 的情况,作母函数

$$F(x) = \sum_{l=0}^{\infty} y(l)x^l. \tag{3}$$

乘以

$$\Phi(x) = \sum_{i=0}^{n} a_ix^i, \tag{4}$$

则得

$$F(x)\Phi(x) = \sum_{l=0}^{\infty}\sum_{i=0}^{n} y(l)a_ix^{i+l}$$

$$= \sum_{p=0}^{\infty} x^p \sum_{i \leqslant \min(n,p)} a_iy(p-i). \tag{5}$$

当 $p \geqslant n$ 时，

$$\sum_{i=0}^{n} a_i y(p-i) = 0.$$

而当 $p < n$ 时，

$$\sum_{i=0}^{p} a_i y(p-i) = \sum_{i=0}^{p} a_i c_{p-i} = b_p,$$

则由 (5) 可知

$$F(x)\Phi(x) = B(x). \tag{6}$$

此处

$$B(x) = \sum_{p=0}^{n-1} b_p x^p,$$

因此

$$F(x) = \frac{B(x)}{\Phi(x)}. \tag{7}$$

求出 $F(x)$ 的幂级数展开式，其 x^l 的系数便是 $y(l)$，即

$$y(l) = \frac{1}{l!} \frac{d^l}{dx^l} F(x) \Big|_{x=0}.$$

这样来求 x^l 的系数有时并不是最方便的，一般可用分项分数法，命

$$\sum_{v=0}^{n} a_v x^v = a_0 \prod_{v=1}^{s} (1 - \lambda_v x)^{l_v}.$$

用分项分数法

$$\frac{B(x)}{\Phi(x)} = \sum_{v=1}^{s} \left(\frac{a_{v1}}{1 - \lambda_v x} + \cdots + \frac{a_v l_v}{(1 - \lambda_v x)^{l_v}} \right),$$

由于

$$\frac{1}{(1 - \lambda_v x)^p} = \sum_{m=0}^{\infty} \frac{p(p+1)\cdots(p+m-1)}{m!} (\lambda_v x)^m,$$

所以

$$y(m) = \sum_{v=1}^{s} \sum_{p=1}^{l_v} a_{vp} \frac{p(p+1)\cdots(p+m-1)}{m!} \lambda_v^m. \tag{8}$$

2) 再看一般的 $g(x)$。在 (5) 中，当 $p \geqslant n$ 时，

$$\sum_{i=0}^{n} a_i y(p-i) = g(p-n).$$

因而

$$F(x)\Phi(x) = B(x) + \sum_{p=n}^{\infty} g(p-n) x^p = B(x) + x^n \sum_{p=0}^{\infty} g(p) x^p.$$

这儿

$$a(x) = \sum_{p=0}^{\infty} g(p) x^p$$

是函数 $g(x)$ 的母函数，因而

$$F(x)=(B(x)+x^n a(x))/\Phi(x). \tag{9}$$

依然可以用分项分数法处理 $1/\Phi(x)$，然后由展开式得 x^l 的系数。

§3. 第二法——降阶法

命

$$\Phi(p)=\sum_{v=0}^{n-1} b_v y(n-1-v+p).$$

作出

$$\begin{aligned}\Phi(p+1)-\lambda\Phi(p)&=\sum_{v=0}^{n-1} b_v y(n-v+p)-\lambda\sum_{v=1}^{n} b_{v-1}y(n-v+p)\\&=\sum_{v=0}^{n}(b_v-\lambda b_{v-1})y(n-v+p)\quad (b_{-1}=0, b_n=0)\\&=\sum_{v=0}^{n} a_v y(n-v+p),\end{aligned}$$

此处

$$a_v=b_v-\lambda b_{v-1}.$$

作多项式

$$\sum_{v=0}^{n} a_v x^v=\sum_{v=0}^{n}(b_v-\lambda b_{v-1})x^v=\sum_{v=0}^{n-1} b_v x^v-\lambda\sum_{v=1}^{n} b_{v-1}x^v=(1-\lambda x)\sum_{v=0}^{n-1} b_v x^v. \tag{1}$$

因此如果 (1) 成立，则差分方程

$$\sum_{v=0}^{n} a_v y(n+p-v)=g(p) \tag{2}$$

可以写成为

$$\Phi(p+1)-\lambda\Phi(p)=g(p),\ \Phi(0)=\sum_{v=0}^{n-1} b_v c_{n-1-v}. \tag{3}$$

形如 (3) 的差分方程已经在 §1 中解决了，解得的 $\Phi(p)$ 再从第 $n-1$ 级的差分方程

$$\sum_{v=0}^{n-1} b_v y(n-1-v+p)=\Phi(p)$$

解决之。

§4. 第三法——Laplace 变换法

解差分方程

$$y(t+1)=ay(t)+g(t), \tag{1}$$

并且要求当 $0\leqslant t\leqslant 1$ 时

$$y(t)=\varphi(t).$$

作 Laplace 变换

$$x(s)=\int_0^\infty e^{-st}y(t)dt,\tag{2}$$

则

$$\int_0^\infty e^{-st}y(t+1)dt=e^s\int_1^\infty e^{-st}y(t)dt=e^s\int_0^\infty e^{-st}y(t)dt-e^s\int_0^1 e^{-st}\varphi(t)dt$$

$$=e^sx(s)-e^s\int_0^1 e^{-st}\varphi(t)dt.$$

另一方面由 (1)

$$\int_0^\infty e^{-st}y(t+1)dt=a\int_0^\infty e^{-st}y(t)dt+\int_0^\infty e^{-st}g(t)dt=ax(s)+\int_0^\infty e^{-st}g(t)dt.$$

因此

$$e^sx(t)-e^s\int_0^1 e^{-st}\varphi(t)dt=ax(s)+\int_0^\infty e^{-st}g(t)dt,$$

即

$$x(s)=\frac{1}{e^s-a}\left[e^s\int_0^1 e^{-st}\varphi(t)dt+\int_0^\infty e^{-st}g(t)dt\right].$$

利用反转公式可得 $y(t)$.

注意　虽然我们讲了三个方法（及以下的第四个方法）但实质上并没有太高的差异，讲来讲去绕不过分项分数这一关．请读者注意．

§5. 第四法——矩阵法

考虑线性差分方程组

$$y_i(t+1)=\sum_{j=1}^n y_j(t)a_{ji}+b_i(t),\quad i=1,2,\cdots,n.\tag{1}$$

n 级差分方程

$$a_0y(t+n)+a_1y(t+n-1)+\cdots+a_ny(t)=g(t)$$

可以看成为 (1) 的特例:

$$\begin{aligned}&y_1(t+1)=y_2(t),\quad && y_1(t)=y(t),\\ &y_2(t+1)=y_3(t) && y_2(t)=y(t+1),\\ &\cdots\cdots\cdots\cdots\cdots && y_3(t)=y(t+2),\\ &y_{n-1}(t+1)=y_n(t) && \cdots\cdots\cdots\cdots\cdots\end{aligned}$$

$$y_n(t+1)=-\frac{1}{a_0}(a_1y_n(t)+\cdots+a_ny_1(t))+\frac{1}{a_0}g(t),$$

把 (1) 式写成为矩阵形式

$$y(t+1)=y(t)A+b(t).\quad y(0)=c.$$

这儿 $y(t)$ 是矢量，$b(t)$ 是已知矢量，A 是方阵，C 是常矢量．

先看 $b(t)=0$ 的情况．

由归纳法易见(比较 §1)

$$y(l)=CA^l.$$

一般的 $b(t)$ 我们可得

$$y(l) = CA^l + \sum_{p=1}^{l} b(p-1)A^{l-p}.$$

写成矩阵表达形式，形式上较复杂的问题一下子变为原则上同样简单的问题了。

§6. 常系数线性微分方程

我们现在考虑微分方程

$$\frac{dx}{dt} = xA + b(t), \quad x(0) = c. \tag{1}$$

此处 x 是未知矢量，($A = A^{(n)}$) 常元素方阵，$b(t)$ 已知矢量.

$$X(t) = e^{At}.$$

已知

$$\frac{dX}{dt} = XA, \quad X(0) = I$$

系以矢量 C，则得

$$\frac{dcX}{dt} = (cX)A, \quad cX(0) = c.$$

因而 $x = cX$ 就是(1)式的解 ($b(t) = 0$).

如果 $b(t) \neq 0$，命

$$x = y(t)e^{At}, \quad y(0) = c,$$

则

$$\frac{dy}{dt}e^{At} + ye^{At}A = ye^{At}A + b(t).$$

即得

$$\frac{dy}{dt} = b(t)e^{-At},$$

因此

$$y(t) - c = \int_0^t b(t)e^{-At}dt,$$

即得

$$x(t) = \left(c + \int_0^t b(\tau)e^{-A\tau}d\tau\right)e^{At}.$$

习题　读者试用

$$\frac{X(t+\delta) - X(t)}{\delta} \text{ 代 } \frac{dX}{dt},$$

而由上节的结果推出本节的结果.

§7. 有重量质点绕地球运动

现在举一个例子，一个有重量的质点在地球表面邻近的真空里，考虑地球运动的作用在内，研究这质点的运动.

命 v 代表质点对地球的速度，命 ω 代表地球的角速度：质点对地球的 coriolis 惯性等于 $2m\omega \times v$（这儿×代表矢量的矢量积，质点对地球的加速度为重力常数 mg。于是质点运动的微分方程式是：

$$\frac{dv}{dt} = g - 2\omega \times v. \tag{1}$$

按矢量积的定义

$$\omega \times v = (\omega_1, \omega_2, \omega_3) \times (v_1, v_2, v_3) = (\omega_2 v_3 - \omega_3 v_2, \omega_3 v_1 - \omega_1 v_3, \omega_1 v_2 - \omega_2 v_1),$$

因此

$$2\omega \times v = -vA, \tag{2}$$

此处

$$A = -2\begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix}.$$

(2) 代入 (1) 式得

$$\frac{dv}{dt} = vA + g, \tag{3}$$

由此解得

$$v = v_0 e^{At} + g\int_0^t e^{A(t-s)} ds = v_0 e^{At} + g\int_0^t e^{A\tau} d\tau. \tag{4}$$

这儿

$$v_0 = v|_{t=0}.$$

再积分，定出动点的动径

$$r = r_0 + v_0\int_0^t e^{A\tau} d\tau + g\int_0^t d\tau\int_0^\tau e^{A\tau_1} d\tau_1^*, \tag{5}$$

此处

$$r_0 = r|_{t=0}.$$

现在求 A^2, A^3, 显然有

$$A^2 = -4(\omega\omega' I - \omega'\omega).$$

又由于 $\omega A = 0$，可知

$$A^3 = -4\omega\omega' A + 4\omega'\omega A = -4\omega\omega' A.$$

因此，当 $l = 0, 1, 2, \cdots$，时，

$$A^{2l+1} = (-4\omega\omega')^l A.$$

当 $l = 1, 2, \cdots$ 时，

$$A^{2l} = (-4\omega\omega')^{l-1} A^2 = (-4\omega\omega')^l I + 4(-4\omega\omega')^{l-1}\omega'\omega.$$

因此

$$e^{At} = I + \sum_{l=0}^{\infty} \frac{(-4\omega\omega')^l t^{2l+1}}{(2l+1)!} A + \sum_{l=1}^{\infty} \frac{t^{2l}}{(2l)!}[(-4\omega\omega')^l I + 4(-4\omega\omega')^{l-1}\omega'\omega]$$

$$= \cos(2\sqrt{\omega\omega'}\,t) I + \frac{\sin(2\sqrt{\omega\omega'}\,t)}{2\sqrt{\omega\omega'}} A - \frac{1}{\omega\omega'}(\cos(2\sqrt{\omega\omega'}\,t) - 1)\omega'\omega.$$

---

* 不要轻易用公式 $\int_0^\tau e^{A\tau} d\tau = (e^{At} - I)A^{-1}$. 因 A^{-1} 不存在。

$$\int_0^t e^{A\tau}d\tau = \frac{\sin 2\sqrt{\omega\omega'}\,t}{2\sqrt{\omega\omega'}}I + \frac{1-\cos 2\sqrt{\omega\omega'}\,t}{4\,\omega\omega'}A$$

$$-\frac{1}{\omega\omega'}\left(\frac{\sin 2\sqrt{\omega\omega'}\,t}{2\sqrt{\omega\omega'}} - t\right)\omega'\omega;$$

$$\int_0^t d\tau\int_0^\tau e^{A\tau_1}d\tau_1 = \frac{1-\cos 2\sqrt{\omega\omega'}\,t}{4\omega\omega'}I + \frac{t - \dfrac{\sin 2\sqrt{\omega\omega'}\,t}{2\sqrt{\omega\omega'}}}{4\omega\omega'}A$$

$$-\frac{1}{\omega\omega'}\left(\frac{1-\cos 2\sqrt{\omega\omega'}\,t}{4\omega\omega'} - \frac{1}{2}t^2\right)\omega'\omega;$$

因此

$$r = r_0 + v_0\left[\frac{\sin 2\sqrt{\omega\omega'}\,t}{2\sqrt{\omega\omega'}}I + \frac{1-\cos 2\sqrt{\omega\omega'}\,t}{4\,\omega\omega'}A\right.$$

$$\left.-\frac{1}{\omega\omega'}\left(\frac{\sin 2\sqrt{\omega\omega'}\,t}{2\sqrt{\omega\omega'}} - t\right)\omega'\omega\right]$$

$$+g\left[\frac{1-\cos 2\sqrt{\omega\omega'}\,t}{4ww}I + \frac{2\sqrt{\omega\omega'}\,t - \sin 2\sqrt{\omega\omega'}\,t}{8\sqrt{\omega\omega'^3}}A\right.$$

$$\left.-\frac{1}{\omega\omega'}\left(\frac{1-\cos 2\sqrt{\omega\omega'}\,t}{4\omega\omega'} - \frac{1}{2}t^2\right)\omega'\omega\right].$$

由于 $xA = -2\omega\times x$. 因此得.

$$r = r_0 + \frac{\sin 2\sqrt{\omega\omega'}\,t}{2\sqrt{\omega\omega'}}v_0 - \frac{1-\cos 2\sqrt{\omega\omega'}\,t}{2\omega\omega'}\omega\times v_0$$

$$-\frac{1}{\omega\omega'}\left(\frac{\sin 2\sqrt{\omega\omega'}\,t}{2\sqrt{\omega\omega'}} - t\right)v_0\cdot\omega'\cdot\omega$$

$$+\frac{1-\cos 2\sqrt{\omega\omega'}\,t}{4\,\omega\omega'}g - \frac{2\sqrt{\omega\omega'}\,t - \sin 2\sqrt{\omega\omega'}\,t}{4(\sqrt{\omega\omega'})^3}\omega\times g$$

$$-\frac{1}{\omega\omega'}\left(\frac{1-\cos 2\sqrt{\omega\omega'}\,t}{4\,\omega\omega'} - \frac{1}{2}t^2\right)g\omega'\omega.$$

当 $v_0 = 0$ 时，

$$r = r_0 + \frac{t^2}{2}g - \frac{2\sqrt{\omega\omega'}\,t - \sin 2\sqrt{\omega\omega'}\,t}{4(\omega\omega')^{3/2}}\omega\times g$$

$$+\frac{1-\cos 2\sqrt{\omega\omega'}\,t - 2\omega\omega' t^2}{4\omega\omega'^{3/2}}\left(-\frac{g\omega'\omega}{\sqrt{\omega\omega'}} + \sqrt{\omega\omega'}\,g\right),$$

第一二项 r_0，$\frac{t^2}{2}g$ 是原位置与由地心引力所得出的，第三项

$$\frac{2\sqrt{\omega\omega'}\,t - \sin 2\sqrt{\omega\omega'}\,t}{4(\omega\omega')^{3/2}}(g\times\omega)$$

表示垂直于子午线平面方向东的偏差．最后一项表示在子午线平面上离开地轴方向（垂直于地轴)的偏差．

对于数值很小的角速度(例如：地球自转 $\sqrt{\omega\omega'}\approx 7.3\times 10^{-5}$秒$^{-1}$)，可以不计$\sqrt{\omega\omega'}$的二次与高次项，因此，由地球所引起的偏差位置有以下的近似公式

$$-\omega\times\left(t^2v_0+\frac{1}{3}t^3g\right).$$

§8. 振　　动

在研究振动现象的时候，常出现二级微分方程组

$$\frac{d^2x}{dt^2}+xA=0. \tag{1}$$

这儿 $x=(x_1,\cdots,x_n)$，$A=A^{(n)}$.

命

$$\frac{dx}{dt}=y,$$

则(1)可以写成为

$$\left(\frac{dx}{dt},\ \frac{dy}{dt}\right)=(x,y)\begin{pmatrix}0 & -A\\ I & 0\end{pmatrix}.$$

由于

$$\begin{pmatrix}I & 0\\ 0 & \sqrt{A}\end{pmatrix}\begin{pmatrix}0 & -A\\ I & 0\end{pmatrix}\begin{pmatrix}I & 0\\ 0 & \sqrt{A}\end{pmatrix}^{-1}=\begin{pmatrix}0 & -\sqrt{A}\\ \sqrt{A} & 0\end{pmatrix}$$

及

$$\begin{pmatrix}I & iI\\ 1 & -iI\end{pmatrix}\begin{pmatrix}0 & -I\\ 1 & 0\end{pmatrix}\begin{pmatrix}1 & iI\\ I & -iI\end{pmatrix}^{-1}=\begin{pmatrix}iI & -I\\ -iI & -I\end{pmatrix}\begin{pmatrix}-iI & -iI\\ -I & 1\end{pmatrix}\frac{1}{-2i}$$

$$=\begin{pmatrix}iI & 0\\ 0 & -iI\end{pmatrix},$$

如果命

$$\Gamma=\begin{pmatrix}I & iI\\ I & iI\end{pmatrix}\begin{pmatrix}I & 0\\ 0 & \sqrt{A}\end{pmatrix}=\begin{pmatrix}I & i\sqrt{A}\\ I & -i\sqrt{A}\end{pmatrix},$$

则

$$\begin{pmatrix}0 & -A\\ I & 0\end{pmatrix}=\Gamma^{-1}\begin{pmatrix}i\sqrt{A} & 0\\ 0 & -i\sqrt{A}\end{pmatrix}\Gamma.$$

即得

$$\left(\frac{dx}{dt},\ \frac{dy}{dt}\right)\Gamma^{-1}=(x,\ y)\Gamma^{-1}\begin{pmatrix}i\sqrt{A} & 0\\ 0 & -i\sqrt{A}\end{pmatrix}.$$

命

$$(x,y)\Gamma^{-1}=(u,v),$$

则

$$\left(\frac{du}{dt},\ \frac{dv}{dt}\right)=(u,v)\begin{pmatrix} i\sqrt{A} & 0\\ 0 & -i\sqrt{A}\end{pmatrix}.$$

即

$$u=ce^{i\sqrt{A}t},\quad v=de^{-i\sqrt{A}t}.$$

因此得出

$$(x,y)=(u,v)\Gamma=(c,d)\begin{pmatrix} e^{i\sqrt{A}t} & 0\\ 0 & e^{-i\sqrt{A}t}\end{pmatrix}\begin{pmatrix} I & i\sqrt{A}\\ I & -i\sqrt{A}\end{pmatrix},$$

即

$$x=ce^{i\sqrt{A}t}+de^{-i\sqrt{A}t},\qquad (2)$$

$$y=i(ce^{i\sqrt{A}t}-de^{-i\sqrt{A}t})\sqrt{A},$$

方程组(1)的初始条件是当 $t=0$ 时,

$$x=x_0,\quad \left(\frac{dx}{dt}\right)_0=v_0.\qquad (=y_0).$$

如此则由(2)可知

$$x_0=c+d,$$
$$v_0=i(c-d)\sqrt{A},\quad c-d=-iv_0A^{-\frac{1}{2}},$$

即

$$c=\frac{1}{2}(x_0-iv_0A^{-\frac{1}{2}}),\quad d=\frac{1}{2}(x_0+iv_0A^{-\frac{1}{2}}).$$

代入(2)得

$$x=\frac{1}{2}(x_0-iv_0A^{-\frac{1}{2}})e^{i\sqrt{A}t}+\frac{1}{2}(x_0+iv_0A^{-\frac{1}{2}})e^{-i\sqrt{A}t}$$

$$=x_0\cos\sqrt{A}\,t+v_0A^{-\frac{1}{2}}\sin\sqrt{A}\,t,\qquad (3)$$

这便是方程(1)的解了.

再考虑非齐次组

$$\frac{d^2x}{dt^2}+xA=f(t)$$

可以写成为

$$\left(\frac{dx}{dt},\ \frac{dy}{dt}\right)=(x,y)\begin{pmatrix} 0 & -A\\ I & 0\end{pmatrix}+(0,f(t)).$$

因而不难算出

$$x=x_0(\cos\sqrt{A}\,t)+v_0(\sqrt{A})^{-1}\sin(\sqrt{A}\,t)$$
$$+\left[\int_0^t f(\tau)\sin(\sqrt{A}(t-\tau))d\tau\right](\sqrt{A})^{-1}.$$

如果把 $t=t_0$ 作为开始时间,则解答换为

$$x=x_0(\cos\sqrt{A}\,(t-t_0))+v_0(\sqrt{A})^{-1}\sin(\sqrt{A}(t-t_0))$$
$$+\left[\int_{t_0}^t f(\tau)\sin(\sqrt{A}(t-\tau))d\tau\right](\sqrt{A})^{-1}.$$

习题取 $f(t)=h\sin(pt+\alpha)$ 的情况.

附注　以上已经看出了方阵函数 e^{At} 的一些用场．但更重要的是它是线性函数方程的研究的核心部分．经过适当的推广，它就成为半群理论的基石．

（参考：E. Hille 与 R. Phillips, Functional Analysis and Semi-groups 1957）。

§9. 矩阵的绝对值

引进矩阵绝对值的方法颇多．我们现在先介绍一下：

命 $A=(a_{ij})1\leqslant i\leqslant m$，$1\leqslant j\leqslant n$，这矩阵的绝对值定义为

$$\|A\|=\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|. \tag{1}$$

不难证明

$$\|A+B\|\leqslant\|A\|+\|B\|, \tag{2}$$

$$\|AB\|\leqslant\|A\|\|B\|. \tag{3}$$

理由是

$$\sum_{i=1}^{m}\sum_{k=1}^{l}\left|\sum_{j=1}^{n}a_{ij}b_{jk}\right|\leqslant\sum_{i=1}^{m}\sum_{k=1}^{l}\sum_{j=1}^{n}|a_{ij}||b_{jk}|\leqslant\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|\cdot\sum_{j_1=1}^{n}\sum_{k=1}^{l}|b_{j_1k}|$$

$$\left\|\int A(t)dt\right\|\leqslant\int\left\|A(t)\right\|dt. \tag{4}$$

如果

$$\sum_{l=0}^{\infty}\|A_l\|<\infty,$$

则 $\sum_{l=0}^{\infty}A_l$ 表 $M\cdot n$ 个绝对收敛级数．

Cauchy 判别条件、考虑矩阵贯

$$\{A_l\}$$

给与 $\varepsilon>0$，存在N使得，p，$q>N$ 时，

$$\|A_p-A_q\|<\varepsilon,$$

则贯 A_l 一定收敛于一个矩阵．

§10. 线性微分方程的唯一存在性问题

我们现在先考虑变系数的线性差分方程

$$x(t+1)=x(t)A(t)+b(t),\quad x(0)=c.$$

算出这个差分方程的解的计算程序是

t	$A(t)$	$b(t)$	$x(t+1)$,
0	$A(0)$	$b(0)$	$x(1)=cA(0)+b(0)$,
1	$A(1)$	$b(1)$	$x(2)=x(1)A(1)+b(1)$,
2	$A(2)$	$b(2)$	$x(3)=x(2)A(2)+b(2)$,

…………………………………………

所以存在性、唯一性问题并没有困难．（因此，一般离散性的问题如果可能，切勿要把它变

为连续性的描述．因为这样做凭空添出不少麻烦).

但是微分方程的情况便不同了．一些性质不太直觉，需要另加处理．

定理 1 假定 $A(t)$ 是一个连续函数 $(t \geqslant 0)$ 方阵．则矢量微分方程

$$\frac{dx}{dt} = xA(t), \quad x(0) = c \tag{1}$$

有唯一的解．这个解当 $t \geqslant 0$ 时存在而且可以写成为

$$x = cX(t) \tag{2}$$

的形式．此处 $X(t)$ 是以下的方阵微分方程的唯一解

$$\frac{dX}{dt} = XA(t), \quad X(0) = I. \tag{3}$$

而

$$\frac{dx}{dt} = xA(t) + g(t) \quad x(0) = c$$

的解答是

$$x = cX(t) + \int_0^t g(\tau)X^{-1}(\tau)d\tau X(t).$$

证明 1) 先用逐次逼近法来证明 (3) 有一解．以积分方程

$$X(t) = I + \int_0^t X(s)A(s)ds \tag{4}$$

来代替微分方程 (3)．

作方阵贯 $\{X_n\}$．

$$X_0 = I,$$
$$X_{l+1} = I + \int_0^t X_l A(s)ds; \quad l = 0, 1, \cdots, \tag{5}$$

故

$$X_{l+1} - X_l = \int_0^t (X_l - X_{l-1})A(s)ds, \quad l = 1, 2\cdots.$$

命

$$m = \max_{0 \leqslant s \leqslant t_1} \|A(s)\|,$$

则当 $0 \leqslant t \leqslant t_1$ 时，

$$\begin{aligned}\|X_{l+1} - X_l\| &= \left\|\int_0^t (X_l - X_{l-1})A(s)ds\right\| \leqslant \int_0^t \|A(s)\|\|X_l - X_{l-1}\|ds \\ &\leqslant m\int_0^t \|X_l - X_{l-1}\|ds.\end{aligned} \tag{6}$$

在此区间内

$$\|X_1 - X_0\| \leqslant \int_0^t \|A(s)\|ds \leqslant mt.$$

用归纳法及关系 (6) 可证：在区间 $0 \leqslant t \leqslant t_1$ 内

$$\|X_{l+1} - X_l\| \leqslant \frac{m^{l+1}t^{l+1}}{(l+1)!},$$

因此级数

$$\sum_{l=0}^{\infty}(X_{l+1}-X_l)$$

在 $0\leqslant t\leqslant x_1$ 中一致收敛，当 $p\to\infty$ 时方阵贯

$$I+\sum_{l=0}^{p}(X_{l+1}-X_l)=X_{p+1}$$

一致收敛于一极限 $X(t)$. 在(5)两边取极限，得出此 $X(t)$ 适合于积分方程(4)，因而也适合于微分方程 (3).

2) 微分方程(1)的存在性；在(3)式左边系以矢量 C，即合所求.

3) 唯一性.

假定 Y 是(3)的另一解，则

$$X-Y=\int_0^t (X(s)-Y(s))A(s)ds. \tag{7}$$

因此

$$\|X-Y\|\leqslant\int_0^t\|A(s)\|\|X(s)-Y(s)\|ds. \tag{8}$$

由于 Y 是可微的，所以 Y 也是连续的.

$$m_1=\max_{0\leqslant t\leqslant t_1}\|X-Y\| \tag{9}$$

是存在的，由(8)可知

$$\|X-Y\|\leqslant m_1\int_0^t\|A(s)\|ds. \tag{10}$$

再把这个估计代入(8)式，得

$$\|X-Y\|\leqslant m_1\int_0^t\|A(s)\|ds\int_0^s\|A(s_1)\|ds_1$$

$$=m_1\iint\limits_{0\leqslant s_1\leqslant s\leqslant t}\|A(s)\|\|A(s_1)\|dsds_1,$$

再代入 (8) 式得

$$\|X-Y\|\leqslant m_1\iiint\limits_{0\leqslant s_2\leqslant s_1\leqslant s\leqslant t}\|A(s)\|\|A(s_1)\|\|A(s_2)\|dsds_1ds_2,$$

等等. 由于

$$\int\limits_{0\leqslant s_l\leqslant s_{l-1}\leqslant\cdots\leqslant s_1\leqslant t}\cdots\cdots\int f(s_1)\cdots f(s_l)ds_l\cdots ds_1$$

$$=\frac{1}{l!}\int_0^t\cdots\int_0^t f(s_1)\cdots f(s_l)ds_1\cdots ds_l=\frac{1}{l!}\left(\int_0^t f(s)ds\right)^l,$$

因此，对任一 $l>0$

$$\|X-Y\|\leqslant\frac{m_1}{l!}\left(\int_0^t\|A(s)\|ds\right)^l.$$

当 $l\to\infty$，右边趋于 0，因此

$$X=Y.$$

(请读者思索一下，这证明的结构在那儿见过没有)

关于非线性部分的解的证明留给读者.

具体写出来

$$X_0 = I,$$

$$X_1 = I + \int_0^t A(s)ds,$$

$$X_2 = I + \int_0^t A(s)ds + \int_0^t A(s)ds\int_0^s A(s_1)ds_1,$$

$$= I + \int_0^t A(s)ds + \iint\limits_{0\leqslant s_1\leqslant s\leqslant t} A(s)A(s_1)dsds_1.$$

(注意由于 $A(s)A(s_1)$ 不一定等于 $A(s_1)A(s)$，所以最后的积分不等于

$$\frac{1}{2}\left(\int_0^t A(s)ds\right)^2.)$$

$$X_3 = I + \int_0^t A(s)ds + \iint\limits_{0\leqslant s_1\leqslant s\leqslant t} A(s)A(s_1)dsds_1$$

$$+ \iiint\limits_{0\leqslant s_2\leqslant s_1\leqslant s\leqslant t} A(s)A(s_1)A(s_2)dsds_1ds_2,$$

因而得出 X 的级数表达法

$$X = I + \int_0^t A(s)ds + \iint\limits_{0\leqslant s_1\leqslant s\leqslant t} (As)A(s_1)dsds_1 + \cdots$$

$$+ \int\limits_{0\leqslant s_l\leqslant\cdots\leqslant s_1\leqslant t}\cdots\cdots\int A(s_1)A(s_2)\cdots A(s_l)ds_1\cdots ds_l + \cdots.$$

如果把条件 $X(0) = I$ 改为 $X(t_0) = I$，即得

$$X = I + \int_{t_0}^t A(s)ds + \iint\limits_{t_0\leqslant s_1\leqslant s\leqslant t} A(s)A(s_1)dsds_1 + \cdots\cdots$$

$$+ \int\limits_{t_0\leqslant s_l\leqslant\cdots\leqslant s_1\leqslant t}\cdots\cdots\int A(s_1)A(s_2)\cdots A(s_l)ds_1\cdots ds_l + \cdots. \tag{11}$$

定义　由(11)所定义的 X 称为函数方阵 $A(t)$ 的由 t_0 到 t 的第积积分，记之为

$$r_{t_0}^t(A) = \widetilde{\int_{t_0}^t}(I + A(t))dt.$$

§11. 第 积 积 分

我们再从差分方程逼迫来看一下第积积分的意义，考虑差分方程

$$X(t+\delta) - X(t) = X(t)\delta A.$$

即

$$X(t+\delta) = X(t)(I + \delta A(t)).$$

因此

$$X(l\delta) = X(0)\prod_{p=0}^{l-1}(I + \delta A(p\delta)).$$

取 $l\delta = t$，则

$$X(t) = X(0)\prod_{p=0}^{l-1}\left(I + \frac{t}{l}A\left(\frac{pt}{l}\right)\right). \quad X(0) = I.$$

因此

$$r_0^t(A(t)) = \widetilde{\int_0^t} (I + A(t))dt = \lim_{l\to\infty} \prod_{p=0}^{l-1} \left(I + \frac{t}{l} A\left(\frac{pt}{l}\right)\right).$$

关于 $r_{t_0}^t$ 有以下诸性质：

(i) $r_{t_0}^t = r_{t_0}^{t_1} r_{t_1}^t$.

证明 $r_{t_0}^{t_1}$ 是

$$\frac{dX}{dt} = XA, \quad X(t_0) = I$$

的解，而 $r_{t_1}^t$ 是

$$\frac{dY}{dt} = YA, \quad Y(t_1) = I$$

的解，命 $XY^{-1} = Z$. 则得

$$\frac{dZ}{dt} = \frac{dX}{dt} Y^{-1} - XY^{-1} \frac{dY}{dt} Y^{-1} = (XA - XA)Y^{-1} = 0,$$

所以 Z 是一常数方阵 C，取 $t = t_1$，即得

$$C = r_{t_0}^{t_1}.$$

(ii) $r_{t_0}^t(A + B) = r_{t_0}^t(P) r_{t_0}^t(A)$.

此处 $P = r_{t_0}^t(A) B (r_{t_0}^t(A))^{-1}$.

证明 $r_{t_0}^t(A + B)$ 及 $r_{t_0}^t(A)$ 各适合于

$$\frac{dX}{dt} = X(A + B), \quad X(t_0) = I.$$

$$\frac{dY}{dt} = YA, \quad Y(t_0) = I.$$

作 $Z = XY^{-1}$，则

$$\frac{dZ}{dt} = \frac{dX}{dt} Y^{-1} - XY^{-1} \frac{dY}{dt} Y^{-1} = [X(A + B) - XA] Y^{-1} = XBY^{-1}$$

$$= Z(YBY^{-1}), \quad Z(t_0) = I.$$

(iii) $r_{t_0}^t = (r_t^{t_0})^{-1}$.

(iv) $r_{t_0}^t(CA(t)C^{-1}) = C r_{t_0}^t(A(t)) C^{-1}$ C 是常数方阵.

(v) 引进第积微商运算

$$D_t X = X^{-1} \frac{dX}{dt}.$$

$\left(\text{由 } \frac{dX}{dt} = XA，\text{可见 } X^{-1}\frac{dX}{dt} = A，\right.$因此 D_t 与 r 是正逆的运算，一个从 X 得 A，而每一从 A 得 $X\Big)$.

显然有

$$D_t(XY) = Y^{-1} D_t X Y + D_t(Y).$$

特例有：

$$D_t(XC) = C^{-1} D_t(X) C,$$

$$D_t(CY) = D_t(Y),$$

$$D_t(X^{-1}) = -XD_t(X)X^{-1}.$$

(vi) $D_t(X') = X'^{-1}(D_t(X)')X'$.

$$\left(X'^{-1}\frac{dX'}{dt} = \left(\frac{dX}{dt}X^{-1}\right)'\right.$$

$$= \left(XX^{-1}\frac{dX}{dt}X^{-1}\right)' = X'^{-1}\left(X^{-1}\frac{dX}{dt}\right)'X'\Big).$$

(vii) 与部分积分相似有公式

$$r_{t_0}^t(Q + D_tX) = X(t_0)^{-1}r_{t_0}^t(XQX^{-1})X(t).$$

证明　左边是

$$\frac{dZ}{dt} = Z\left(Q + X^{-1}\frac{dX}{dt}\right), \quad Z(t_0) = I$$

的解答，命 $Y = ZX^{-1}$，则得

$$\frac{dY}{dt} = \frac{dZ}{dt}X^{-1} - ZX^{-1}\frac{dX}{dt}X^{-1} = Z\left(Q + X^{-1}\frac{dX}{dt}\right)X^{-1} - ZX^{-1}\frac{dX}{dt}X^{-1}$$

$$= ZQX^{-1} = YXQX^{-1}, \quad Y(t_0) = X(t_0)^{-1}.$$

因此，

$$Y(t) = X(t_0)^{-1}r_{t_0}^t(XQX^{-1}).$$

§12. 解的满秩性

定理 1　假定 $\int_0^{t_1}\|A(t)\|dt$ 存在，则在 $0 \leqslant t \leqslant t_1$ 内方程

$$\frac{dX}{dt} = XA(t), \quad X(0) = I \tag{1}$$

的解 $X(t)$ 是满秩的.

如果 $A(t) = A$ 是常数方阵，则由 §6 已知方程 (1) 的解等于

$$e^{At}$$

在上一章中已经证过它是满秩的了.

这定理是以下的 Jacobi 恒等式的推论：

定理 2　$X(t)$ 的行列式

$$|X(t)| = e^{\int_0^t tr(A(s))ds}.$$

证明　由行列式的微分性质

$$\frac{d}{dt}\begin{vmatrix} x_{11}(t), & \cdots, & x_{1n}(t) \\ x_{21}(t), & \cdots, & x_{2n}(t) \\ \cdots & \cdots & \cdots \\ x_{n1}(t), & \cdots, & x_{nn}(t) \end{vmatrix} = \begin{vmatrix} \frac{dx_{11}}{dt}, & x_{12}, & \cdots, & x_{1n} \\ \frac{dx_{21}}{dt}, & x_{22}, & \cdots, & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{dx_{n1}}{dt}, & x_{n2}, & \cdots, & x_{nn} \end{vmatrix}$$

$$
+\begin{vmatrix} x_{11}, & \frac{dx_{12}}{dt}, & \cdots, & x_{1n} \\ x_{21}, & \frac{dt_{22}}{dt}, & \cdots, & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1}, & \frac{dx_{n2}}{dt}, & \cdots, & x_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} x_{11}, & x_{12}, & \cdots, & \frac{dx_{1n}}{dt} \\ x_{21}, & x_{22}, & \cdots, & \frac{dx_{2n}}{dt} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1}, & x_{n2}, & \cdots, & \frac{dx_{nn}}{dt} \end{vmatrix}.
$$

看其中的一个，例如第一个，因为

$$
\frac{dx_{i1}}{dt} = \sum_{j=1}^{n} x_{ij} a_{j1},
$$

所以

$$
\begin{vmatrix} \frac{dx_{11}}{dt}, & x_{12}, & \cdots, & x_{1n} \\ \frac{dx_{21}}{dt}, & x_{22}, & \cdots, & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{dx_{n1}}{dt}, & x_{n2}, & \cdots, & x_{nn} \end{vmatrix} = \begin{vmatrix} \sum_{j=1}^{n} x_{1j} a_{j1}, & x_{12}, & \cdots, & x_{1n} \\ \sum_{j=1}^{n} x_{2j} a_{j1}, & x_{22}, & \cdots, & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{j=1}^{n} x_{nj} x_{j1}, & x_{n2}, & \cdots, & x_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} x_{11}, & \cdots, & x_{1n} \\ & & \\ \cdots & \cdots & \cdots \\ x_{n1}, & \cdots, & x_{nn} \end{vmatrix},
$$

因此

$$
\frac{d}{dt}|X(t)| = (a_{11} + \cdots + a_{nn})|X(t)|.
$$

即得定理.

显然对任一常数方阵 C，$Y = CX$ 适合于

$$
\frac{dY}{dt} = YA(t). \tag{2}
$$

因此 CX 是方程 (2) 的通解. 反之，如果 Y 适合于 (2)，而 $Y(0) = C$，把 Y, C 写成矢量形式

$$
\begin{pmatrix} y^{(1)} \\ \vdots \\ y^{(n)} \end{pmatrix}, \quad \begin{pmatrix} C^{(1)} \\ \vdots \\ C^{(n)} \end{pmatrix},
$$

则

$$
\frac{dy^{(i)}}{dt} = y^{(i)} A(t), \quad y^{(i)}(0) = C^{(i)}.
$$

因此

$$
y^{(i)} = C^{(i)} X.
$$

而

$$
Y = CX.
$$

因此 Y 的秩等于初始值 C 的秩.

§13. 非齐次方程

考虑

$$\frac{dx}{dt}=xA(t)+f(t),\quad x(0)=C. \tag{1}$$

我们还是用“参数变动术”. 命

$$x=y(t)X, \tag{2}$$

此处X是

$$\frac{dX}{dt}=XA(t)\quad X(0)=I \tag{3}$$

的解,把(2)代入(1)得

$$\frac{dy}{dt}X+y\frac{dX}{dt}=yXA+f(t),$$

即得

$$\frac{dy}{dt}X=f(t).$$

因此

$$\frac{dy}{dt}=f(t)X^{-1},$$

或

$$y=C+\int_0^t f(s)X^{-1}(s)ds.$$

由此推得(1)的解法公式

$$x=CX(t)+\left(\int_0^t f(s)X(s)^{-1}ds\right)X(t). \tag{4}$$

附记 1. 方程(3)可以写成为

$$X^{-1}\frac{dX}{dt}=A(t),\quad X(0)=I.$$

由 $\frac{dX^{-1}}{dt}=-X^{-1}\frac{dX}{dt}X^{-1}$ 可知

$$\frac{dX^{-1}}{dt}X=-A(t).$$

如命 $X^{-1}=Y$, 则Y适合于

$$\frac{dY}{dt}=-A(t)Y,\quad Y(0)=I. \tag{5}$$

而解答(4)式可以写成为

$$x=\left[C+\int_0^t f(s)Y(s)ds\right]Y(t)^{-1}. \tag{6}$$

2. 如果把(3)的初始条件改为

$$X(t_0)=I,$$

命 X_1 是适合于

$$\frac{dX_1}{dt} = X_1A(t), \quad X_1(t_0) = I$$

的解，则由于

$$CX(t), \quad (C \text{ 任一满秩方程})$$

是 (3) 式的通解，所以

$$X_1(t) = CX(t), \quad X_1(t_0) = CX(t_0).$$

即得

$$X_1(t) = X(t_0)^{-1}X(t).$$

§ 14. 微 扰 理 论

非齐次方程的另一应用是微扰理论，我们要把

$$e^{A+\varepsilon B}$$

展开为 ε 的幂级数

$$e^{A+\varepsilon B} = e^A + \sum_{n=1}^{\infty} \varepsilon^n Q_n(A, B).$$

当 $AB = BA$ 时问题不大，但当 $AB \neq BA$ 时情况变得复杂得多，我们现在用非齐次微分方程的解法来处理这一问题.

考虑

$$\frac{dX}{dt} = X(A + \varepsilon B), \quad X(0) = I. \tag{1}$$

这方程式的解是

$$e^{(A+\varepsilon B)t}.$$

由

$$\frac{d(Xe^{-At})}{dt} = \frac{dX}{dt}e^{-At} - XAe^{-At} = \varepsilon XBe^{-At}$$

可知

$$X(t) = \left(I + \varepsilon\int_0^t XBe^{-As}ds\right)e^{At} = e^{At} + \varepsilon\int_0^t X(s)Be^{A(t-s)}ds. \tag{2}$$

这是一个迭代公式，例如把积分号下的 $X(s)$ 取以 (2) 代入，则得

$$X(t) = e^{At} + \varepsilon\int_0^t\left[e^{As} + \varepsilon\int_0^s X(s_1)Be^{A(s-s_1)}ds_1\right]Be^{A(t-s)}ds$$

$$= e^{At} + \varepsilon\int_0^t e^{As}Be^{A(t-s)}ds + \varepsilon^2\int_0^t ds\int_0^s X(s_1)Be^{A(s-s_1)}Be^{A(t-s)}ds_1$$

还可再迭代得出 ε^3, ε^4, $\cdots$ 等项来.

取 $t = 1$，得出

$$e^{A+\varepsilon B} = e^A + \varepsilon\int_0^1 e^{As}Be^{A(1-s)}ds + \varepsilon^2\int_0^1 ds\int_0^s X(s_1)Be^{A(s-s_1)}Be^{A(1-s)}ds_1.$$

附记　读者想一下，这方法和上节的方法有没有共同之点.

§15. 函 数 方 程

命

$$Y(t)=e^{At},$$

它适合于函数方程

$$Y(s+t)=Y(s)Y(t),\quad -\infty<s,\ t<\infty. \tag{1}$$

问题在于有没有其它的矩阵函数仍然适合于(1)，解答是肯定的，由于

$$Y(t)=\begin{pmatrix}e^{at} & 0\\ 0 & 0\end{pmatrix}$$

是一解．但当添一个要求：有一个数值 t_0 使 $Y(t_0)$ 是满秩的，则解答是否定的．由

$$Y(t_0)=Y(0)Y(t_0),$$

因此

$$Y(0)=I.$$

由于

$$Y(t)Y(-t)=Y(0)=I,$$

所以所有的 $Y(t)$ 都是满秩的．

如果 $Y(t)$ 可微分，问题极易解决，(1)式对 s 及对 t 微商即得

$$Y'(s+t)=Y'(s)Y(t)$$
$$Y'(s+t)=Y(s)Y'(t).$$

因此

$$Y'(s)Y(t)=Y(s)Y'(t).$$

即对所有的 s，t

$$Y^{-1}(s)Y'(s)=Y'(t)Y^{-1}(t).$$

因此 $Y^{-1}(s)Y'(s)$ 是一个常数方阵 A，即

$$\frac{d}{ds}Y(s)=Y(s)A.\ \ Y(0)=I,$$

解得

$$Y(t)=e^{At}.$$

我们可以减弱假定：

定理 1 假定 (i) 当 $0\leqslant s,t,s+t\leqslant t_0$ 时有

$$Y(s+t)=Y(s)Y(t).$$

(ii) 在 $0\leqslant t\leqslant t_0$ 中 $Y(t)$ 是连续函数及 (iii) 至少有一个 t_1，使 $Y(t_1)$ 是满秩的，在这些假定下

$$Y(t)=e^{At}.$$

证明 把积分变为和的极限

$$\int_0^t Y(s)ds=\lim_{\delta\to 0}\sum_{k=0}^{N-1}Y(k\delta)\delta,\qquad \delta=\frac{t}{N}.$$

$$= \lim_{\delta \to 0} \sum_{k=0}^{N-1} Y(\delta)^k \delta,$$

$$= \lim_{\delta \to 0} \frac{Y(N\delta) - I}{Y(\delta) - I} \delta = \lim_{\delta \to 0} \frac{\delta}{Y(\delta) - I} (Y(t) - I). \qquad (2)$$

由连续性可知

$$Y(t) = I + o(1),$$

因此

$$\int_0^t Y(s) ds = tI + o(t).$$

所以当 t 充分小时，

$$\int_0^t Y(s) ds$$

是满秩，因此对固定的足够小的 t，当 δ 充分小时，

$$\sum_{k=0}^{N-1} Y(\delta)^k \delta$$

也是满秩的.

当 $\delta \to 0$ 时，

$$\frac{Y(\delta) - I}{\delta} \sum_{k=0}^{N-1} Y(k\delta)\delta = Y(t) - I. \qquad (3)$$

由于左边第二因子是满秩的，因此 (1) 与 (2) 得

$$A = \lim_{\delta \to 0} \frac{Y(\delta) - I}{\delta} = (Y(t) - I) \left(\int_0^t Y(s) ds \right)^{-1}.$$

即得

$$A \int_0^t Y(s) ds = Y(t) - I.$$

因为左边可微分，因此 $Y(t)$ 是可微分的，而且对充分小的 t，

$$AY(t) = Y'(t), \quad Y(0) = I.$$

即对充分小的 t 有

$$Y(t) = e^{At}.$$

由于

$$Y(nt) = (Y(t))^n,$$

因而对大 t 也对.

§ 16. 解微分方程 $\frac{dX}{dt} = AX + XB$

我们现在考虑较复杂的微分方程

$$\frac{dX}{dt} = A(t)X + XB(t), \quad X(0) = C. \qquad (1)$$

这个微分方程形式上可能较复杂些，但是如果用 X 的元素 $x_{rs}(t)$ 写出来，定依然是 n^2 个变元的 n^2 个方程组，因而解法是存在的，是唯一的，问题在于具体地解出这个问题.

先从

$$\frac{dY}{dt}=A(t)Y,\qquad Y(0)=I$$

及

$$\frac{dZ}{dt}=ZB(t),\qquad Z(0)=I$$

出发，由于

$$\frac{d(YCZ)}{dt}=\frac{dY}{dt}CZ+YC\frac{dZ}{dt}=A(t)YCZ+YCZB(t)$$

及 $(YCZ)_{t=0}=C$，因此

$$X=YCZ$$

就是(1)式的解——唯一解.

特例有

定理 1 对常数方阵 A 与 B，微分方程

$$\frac{dX}{dt}=AX+XB,\quad X(0)=C \tag{2}$$

的解是

$$X=e^{At}Ce^{Bt}. \tag{3}$$

这是一个简单的微分方程，但在量子力学上有用，其解的性质在量子力学上有很自然的意义（例如见 D. Haas, *Elements of statistical Mechanics*, 1954, P. 149.）.

这类微分方程在 Лаппо-Аанилевссит 的专著上作了详尽的讨论 (Труды 1-гс Всесоюзн. сбезда матем. 1936).

这定理有以下的应用:

定理 2 如果

$$X=-\int_0^{\infty}e^{At}Ce^{Bt}dt \tag{4}$$

对所有的 C 都存在，则矩阵方程

$$AX+XB=C \tag{5}$$

有唯一的解.

证明 先考虑方程

$$\frac{dZ}{dt}=AZ+ZB,\quad Z(0)=C,$$

两边由 0 到∞对 t 积分，并且假定

$$\lim_{t\to\infty}Z(t)=0, \tag{6}$$

如此则得

$$-C=-Z(0)=A\int_0^{\infty}Z(t)dt+\int_0^{\infty}Z(t)dtB.$$

因此，

$$X=-\int_0^{\infty}Z(t)dt=-\int_0^{\infty}e^{At}ce^{Bt}dt$$

就是(5)的一个解，由于条件(4)，可知条件(6)是适合的.

关于唯一性，我们可以把(5)看成为 n^2 个同变数的 n^2 个线性方程组，既然对所有的 C 都有解，则这 n^2 个线性方程组的行列式 $\neq 0$，因此，对每一 C 有唯一的解.

附记　在什么条件下

$$\int_0^{\infty} e^{At} c e^{Bt} dt$$

对所有的 C 都存在，将来将证明：方程组(5)有解的必需且充分的条件是 A 的特征根 λ_i 与 B 的特征根 μ_j 的和 $\lambda_i + \mu_j$ 都不等于 0.

第五章　解的渐近性质

§1. 常系数差分方程

还是从差分方程的研究开始；已知常系数的差分方程

$$x(t+1)=x(t)A,\quad x(0)=c \tag{1}$$

的解是

$$x(t)=cA^t,\quad t=0,1,2,3,\cdots. \tag{2}$$

现在我们要研究的问题是：当 $t\to\infty$ 时，解 $x(t)$ 的性质，当 $t=0,1,2,\cdots$ 时

$$x(t)=(x_1(t),\cdots,x_n(t))$$

代表 n 维空间的一个贯．我们可以问：这贯是否有界？是否有极限？是否会回到原出发处？是否会"遍历"一条曲线上所有的点？还是否"对一切 c" 都有某种性质，还是对"一些 c" 或" 那些 c" 有某种性质？一切的一切归结为方阵方幂

$$A^t,\quad t=0,1,2,\cdots$$

的研究，这是我们第三章中所研究过的问题，在讨论之前我们先研究一下，当 α 是实数时，

$$\xi^t=e^{2\pi i\alpha t},\quad t=0,1,2,\cdots$$

的性质，这些点在复平面的单位圆上，首先如果 α 是有理数 $p/q(q>0,p,q$ 互素)，则点列

$$\xi^t$$

中只有有限个值 $t=0,\cdots,q-1$，其后便周而复始地循迴往复，如果 α 是无理数，ξ^t 在单位圆上成一无处不稠密的点集，即对圆上任一点 $\xi_0=e^{2\pi i\alpha_0}(0\leqslant\alpha_0<1)$，给任一 $\varepsilon>0$，我们有无数个 t 使

$$|\xi_0-\xi^t|<\varepsilon$$

(要证明这个事实，需要些数论知识，即所谓 Kronecker 定理：给了任一 $\varepsilon>0$，有无数个自然数 q 及整数 p 使

$$|\alpha p-\alpha_0-q|<\varepsilon.$$

如果这个对了，则

$$|\xi^p-\xi_0|<2\pi\varepsilon.$$

即得所证了)．

这样的性质称为点列 ξ^t 遍历圆上各点，现在回来研究

$$A^t,\quad t=0,1,2,\cdots$$

有一个 P 使 PAP^{-1} 等于 Jordan 块

$$J_\lambda=\begin{cases}\lambda J & \text{当 } \lambda\neq 0,\ J=\begin{pmatrix}110 & \\ 011 & \\ 001\cdot & \cdot\end{pmatrix}\\ \\ J_0 & \text{当 } \lambda=0,\ J_0=\begin{pmatrix}010\cdots\\ 001\cdots\\ 000\cdots\\ \cdots\cdots\\ \cdots\cdots\end{pmatrix}.\end{cases}$$

如果 $|\lambda|<1$，则

$$\lim_{t\to\infty}J_\lambda^t=0.$$

如果 $|\lambda|>1$，则

$$\lim_{t\to\infty}J_\lambda^t$$

不存在.

如果 $|\lambda|=1$，$\lambda=e^{2\pi i\alpha}$，而 J_λ 非单构的，则极限依然不存在.

如果 $|\lambda|=1$，而且 J_λ 是单构的，则问题化为以上所讨论的结果.

把方阵 PAP^{-1} 折为三部分：

(i) 特征根绝对值<1 的诸子块留下，其他的部分记之为 0. 这样的方阵用 B_-表之.

(ii) 特征根绝对值$=1$，而且是单构的诸子块留下，其它的记之为 0，这样的方阵用 B_0 表之.

(iii) 以上未用到的子块留下，其它的记之为 0，这样的方阵用 B_+ 表之.

这样

$$PAP^{-1}=B_-+B_0+B_+.$$

其中任二之积等于 0. 并且 $\lim B_-^t$ 存在，$\lim B^t$. 有界，$\lim B_-^t$ 无界，命

$$\begin{aligned}A&=P^{-1}(B_-+B_0+B_+)P\\&=A_-+A_0+A_+,\end{aligned}$$

这样便把 A 分为三份了.

现在回过去考虑原问题：

$$x=cA^t=c(A_-^t+A_0^t+A_+^t).$$

如果 $cA_+=0$，则

$$\lim_{t\to\infty}x(t)$$

有界，如果 $cA_+=0$ 而且 $cA_0=0$，则

$$\lim_{t\to\infty}x(t)=0.$$

由此推出几个简单的结论：

定理 1 如果 A 的特征值的绝对值都小于 1，则不管怎样的初始值，

$$\lim_{t\to\infty}x(t)=0.$$

这样的性质称为“渐近稳定性”.

定理 2 如果 A 的特征值的绝对值$\leqslant 1$，而绝对值等于 1 的部分是单构的，则 $x(t)$ 是

有界的．并且给一 $\varepsilon>0$，我们能找到 δ 使

$$\|c-c_1\|<\delta$$

时，

$$\|x(t)-x_1(t)\|<\varepsilon,$$

这儿 $x_1(t)=c_1A^t$.

定理 3 如果 A 有一个特征根等于 1，而其他的都小于 1，则 $\lim\limits_{t\to\infty}x(t)$ 存在，其极限的支量间的比例不变．

证明 这样 $\lim\limits_{t\to\infty}A^t$ 的秩等于 1，即

$$\lim_{t\to\infty}A^t=u'v,\quad uv'=1,$$

这儿 u,v 是两个矢量，因此

$$\lim_{t\to\infty}x(t)=\lim_{t\to\infty}cA^t=(cu')v.$$

即与 v 成比例，即得所证．

读者试研究 A 的特征根的绝对值都 $=1$，而且是单构的情况．

看来变化多端，实质脱胎于一 $(n=1)$，标准法式稳抓住，便可推出一切．

§2. 广相似性

现在考虑较一般的差分方程组

$$x(t+1)=x(t)A(t). \tag{1}$$

再考虑换变数的情况：命

$$y(t)=x(t)L(t). \tag{2}$$

则得

$$\begin{aligned}y(t+1)&=x(t+1)L(t+1)\\&=x(t)A(t)L(t+1)\\&=y(t)L(t)^{-1}A(t)L(t+1).\end{aligned} \tag{3}$$

这样便把以 $A(t)$ 为系数方阵的差分方程组变为以

$$B(t)=L(t)^{-1}A(t)L(t+1) \tag{4}$$

为系数方阵的差分方程组了．

定义 1 如果 $L(t)$ 与 $L(t)^{-1}$ 都有界，则 $L(t)$ 称为 Ляпунов 方阵．

如果 $L(t)$ 是 Ляпунов 方阵，则由

$$\lim_{t\to\infty}x(t)=0$$

可推得

$$\lim_{t\to\infty}y(t)=0$$

并且反之亦真．

定义 1 如果 $A(t)$ 与 $B(t)$ 适合于(4)而且 $L(t)$ 是 Ляпунов 方阵，则此二方阵称为 л 相似性，记之为

$$A\overset{\text{л}}{=}B.$$

显然有(i), $A \stackrel{л}{=} A$ (ii) 如果 $A \stackrel{л}{=} B$ 则 $B \stackrel{л}{=} A$; (iii) 如果 $A \stackrel{л}{=} B$, $B \stackrel{л}{=} C$, 则 $A \stackrel{л}{=} C$.

又如果有常数方阵 P 使

$$PAP^{-1} = B,$$

则 $A \stackrel{л}{=} B$.

定义 2　凡 л 相似于常数方阵的方阵称为可化方阵.

定理 1　可化方阵一定 л 相似于一个特征根是实数的方阵.

证明　由于可化方阵一定 л 相似于一个常数方阵因此我们只研究常数方阵，还是从考虑 $n = 1$ 时的情况出发

$$x(t+1) = x(t)\rho e^{i\theta}.$$

命 $y(t) = x(t)l^{-1}(t)$, 则

$$y(t+1) = y(t)l(t)\rho e^{i\theta}/l(t+1).$$

取

$$l(t) = e^{i\theta t}$$

即可使

$$y(t+1) = y(t)\rho.$$

一般的情况可由标准型推出.

§ 3. 常数系数线性常微分方程组

我们现在研究常微分方程的解的渐近性质：方程组

$$\frac{dx}{dt} = xA, \qquad x(0) = c \tag{1}$$

的解

$$x(t) = ce^{At}. \tag{2}$$

如前，我们只需要研究，当 $t \to \infty$ 时，

$$e^{At}$$

的性质.

也是把 A 化成为 Jordan 标准型来讨论.

(i)　如果 λ 的实数部分都是负的，则

$$\lim_{t\to\infty} e^{\lambda/t} = 0.$$

(ii)　如果 λ 是纯虚的而且是单构的，则

$$\lim_{t\to\infty} e^{i\lambda t}$$

不存在 ($\lambda = 0$ 除外)，而 $z = e^{i\lambda t}$ 在平面上画一圆圈.

(iii) 不然，则

$$x(t) \qquad 0 \leqslant t \leqslant \infty$$

画出一条通过 $x = c$ 的曲线，而且趋向∞.

因此也推得

定理 1　如果 A 的特征根的实数部分都为负，则不管初始值怎样，积分曲线 $(x_1(t), \cdots x_n(t))$ 趋向原点.

定理 2 如果 A 是单构的，而且特征根是纯虚的，则积分曲线在空间画一闭曲线.

证明 有 P 使

$$PAP^{-1}=\begin{pmatrix}0 & \lambda_1\\ -\lambda_1 & 0\end{pmatrix}\dotplus\cdots\dotplus\begin{pmatrix}0 & \lambda_V\\ -\lambda_V & 0\end{pmatrix}\dotplus O\dotplus\cdots\dotplus O.$$

命

$$xP^{-1}=y,\ \text{则}$$

$$\frac{dy}{dt}=y(PAP^{-1}).$$

因为

$$e^{\begin{pmatrix}0 & \lambda\\ -\lambda & 0\end{pmatrix}t}=\sum_{n=0}^{\infty}\frac{1}{n!}\begin{pmatrix}0 & \lambda\\ -\lambda & 0\end{pmatrix}^n t^n$$

$$=\sum_{l=0}^{\infty}\frac{1}{(2l)!}\begin{pmatrix}(-\lambda^2t^2)^l & O\\ 0 & (-\lambda^2t^2)^l\end{pmatrix}+\sum_{l=0}^{\infty}\frac{(-\lambda^2t^2)^l}{(2l+1)!}\begin{pmatrix}0 & \lambda\\ -\lambda & 0\end{pmatrix}t$$

$$=\begin{pmatrix}\cos\lambda t & \sin\lambda t\\ -\sin\lambda t & \cos\lambda t\end{pmatrix}.$$

因此

$$y=c\left(\begin{pmatrix}\cos\lambda_1 t & \sin\lambda_1 t\\ -\sin\lambda_1 t & \cos\lambda_1 t\end{pmatrix}\dotplus\cdots\right)$$

$$y_1=c_1\cos\lambda_1 t-c_2\sin\lambda_1 t$$

$$y_2=c_1\sin\lambda_1 t+c_2\cos\lambda_1 t,\ \text{等}.$$

而

$$y_1^2+y_2^2=c_1^2+c_2^2, y_3^2+y_4^2=c_3^2+c_4^2,\cdots.$$

这说明了，积分曲线是闭曲线.

附记 1

$$yy'=cc'.$$

所以

$$xP^{-1}P'^{-1}x'=cc'.$$

这说明积分曲线在一相似的椭球上，实质上远不止这一性质.

附记 2 我们所考虑的积分曲线都是从 $x(0)=c(\neq 0)$ 出发，如果 $c=0$，则由上可知通过原点可能不止一条积分曲线，或者没有积分曲线，这样的点称为奇点.

习题 1 试用特征根分类来研究原点附近的二个变数的方程组的情况.
(已经学过了)

习题 2 试研究三个变数三个方程式的情况.

习题 3 解方程组

$$\frac{dx}{dt}=xA+b. \qquad x(0)=c$$

此处 b 是常数矢量.

§4. Ляпунов 法介绍

以下所介绍的 Ляпунов 法可以广泛地用来研究非线性函数方程的解的稳定性，但这

儿可以最清楚地看出其实质.

还是研究

$$\frac{dx}{dt}=xA,\quad x(0)=c. \tag{1}$$

这儿 c 与 A 是实的. 命

$$u=xYx'$$

这儿 Y 是一个待定的常数对称方阵,微分得

$$\frac{du}{dt}=\frac{dx}{dt}Yx'+xY\left(\frac{dx}{dt}\right)'=x(AY+YA')x'.$$

假定我们能够选得一个定正的 Y 使

$$AY+YA'=-I,$$

则

$$\frac{du}{dt}=-xx'\leqslant-\lambda_0^{-1}u.$$

这儿 λ_0 是 Y 的最大的特征根,由于 $u(0)$ 是正的, 因此

$$d\log u(t)=\frac{du}{u}\leqslant-\lambda_0^{-1}dt,$$

$$\int_0^t d\log u(t)\leqslant-\lambda_0 t,$$

因此

$$u(t)\leqslant u(0)e^{-\lambda_0^{-1}t}.$$

当 $t\to\infty$ 时,得

$$u=xYx'\to 0.$$

由于 Y 的定正性,因此 $x\to 0$.

因此研究

$$\lim_{t\to\infty}x(t)$$

是否趋于 0 的问题一变而为 Y 的存在性的问题.

1) 如果 A 的特征根的实数部分都<0, 则 Y 一定存在.

这个 Y 就是

$$Y=\int_0^\infty e^{At}e^{A't}dt.$$

这是对称的,收敛的,而且是定正的,前二性质十分容易看出,今看第三性质: 如果

$$0=xYx'=\int_0^\infty(xe^{At})(xe^{At})'dt.$$

因此得出 $xe^{At}=0$. 但指数方阵 e^{At} 是满秩的,因此 $x=0$, 所以 Y 是定正的.

又由部分积分法可知

$$AY+YA'=\int_0^\infty(Ae^{At}\cdot e^{A't}+e^{At}\cdot e^{A't}A')dt$$
$$=\int_0^\infty\frac{d}{dt}(e^{At}\cdot e^{A't})dt=-1.$$

2）如果有定正的对称的 Y 而且使

$$AY + YA' = -I,$$

则 A 的特征根的实数部分都是负的.

证明　命 ρ 是 A 的特征根，v 是对应的矢量. 即

$$vA = \rho v.$$

由于 A 是实的，所以

$$\bar{v}A = \overline{\rho v}.$$

在 $AY + YA' = -I$ 左右各乘以 $v, \bar{v}'$ 得

$$v(AY + YA')\bar{v}' = -v\bar{v}'$$

即得

$$(\rho + \bar{\rho})vY\bar{v}' = -v\bar{v}'.$$

因此

$$\rho + \bar{\rho} < 0.$$

即得所证.

我们用 Ляпунов 法处理一个较广泛的问题：

$$\frac{dx}{dt} = x(A + B(t)),$$

这儿 A 仍然有以上的性质，但

$$\lim_{t\to\infty} \|B(t)\| = 0.$$

我们仍然用以上的 Y；则

$$\begin{aligned}\frac{d}{dt}(xYx') &= \frac{dx}{dt}Yx' + xY\left(\frac{dx}{dt}\right)' \\ &= x(A + B(t))Yx' + xY(A' + B'(t))x' \\ &= -xx' + x(B(t)Y + YB'(t))x'.\end{aligned}$$

当 t 充分大时，($t \geqslant t_0$ 时)

$$\frac{d}{dt}(xYx') \leqslant -\lambda_1 xYx'.$$

此处 λ_1 是一常数 > 0，积分得

$$xYx' \leqslant (xYx')_{t=t_0} e^{-\lambda_1 t}.$$

因此当 $t \to \infty$ 时 $xYx' \to 0$，而 $x \to 0$.

附记　如果 A 不是常数方阵，我们不能得出同样的结果来，也就是有 $A(t)$ 存在.

$$\frac{dy}{dt} = yA(t)$$

的解适合于 $\lim\limits_{t\to\infty} y = 0$，而

$$\frac{dx}{dt} = x(A(t) + B(t))$$

的解却无此性质，虽然我们假定了

$$\lim_{t\to\infty}\|B(t)\| = 0.$$

反例是:

$$\frac{dy_1}{dt} = -ay_1, \qquad \frac{dy_2}{dt} = (\sin\log t + \cos\log t - 2a)y_2$$

的解是

$$y_1 = c_1 e^{-at}, \quad y_2 = c_2 e^{t\sin\log t - 2at}.$$

如果 $a > \frac{1}{2}$，当 $t\to\infty$ 时，都趋于 0. 如果添上

$$B(t) = \begin{pmatrix} 0 & e^{-at} \\ 0 & 0 \end{pmatrix},$$

则新方程组是

$$\frac{dx_1}{dt} = -ax_1,$$

$$\frac{dx_2}{dt} = (\sin\log t + \cos\log t - 2a)x_2 + x_1 e^{-at},$$

而解答是

$$x_1 = c_1 e^{-at},$$

$$x_2 = e^{t\sin\log t - 2at}\left(c_2 + c_1\int_0^t e^{-\tau\sin\log\tau}d\tau\right).$$

考虑

$$t = e^{\left(2n+\frac{1}{2}\right)\pi}, \qquad n = 1, 2\cdots$$

诸点，则

$$\int_0^t e^{-\tau\sin\log\tau}d\tau > \int_{te^{-\pi}}^{te^{-2\pi/3}} e^{-\tau\sin\log\tau}d\tau$$

$$= \pi\int_{2/3}^1 \exp(-te^{-\alpha\pi}\cos\alpha\pi)te^{-\alpha\pi}d\alpha \ (\text{取}\ \tau = te^{-\alpha\pi})$$

$$> \pi t\exp(te^{-\pi}/2)\int_{2/3}^1 e^{-\alpha\pi}d\alpha = t(e^{-2\pi/3} - e^{-\pi})\exp(te^{-\pi}/2).$$

所以在 x_2 的解答中 c_1 的系数

$$> e^{t(1-2a)}t(e^{-2\pi/3} - e^{-\pi})e^{\frac{1}{2}te^{-\pi}}.$$

当

$$1 - 2a > -\frac{1}{2}e^{-\pi}$$

时，$\lim_{t\to\infty} x_n$ 或趋于∞，或等于 0，或趋于$-\infty$，视 $c_1 > 0$，$= 0$ 或 < 0 而定. 因此，我们不能得出我们所希望的结论.

习题 1　如果 A 是复元素方阵，我们怎样来运用 Ляпунов 方法.

§5. 稳　定　性

一个物理系统有时为一微分方程组

$$\frac{dx_i}{dt}=f_i(x_1,\cdots,x_n,t),\quad i=1,2,\cdots,n \tag{1}$$

所刻划，其中 $x_1,\cdots,x_n$ 是参数，而 t 是时间．研究这物理系统在平衡状态附近的性质是一个十分重要的问题．即如果来一个微小干扰，这个系统还会还原的则称为稳定的．不然，则称为不稳定的．当然物理系统可以通过试验而辨别其是否稳定，但有时实验花费大，需时久，不如用些数学方法来判断其是否稳定．

确切的稳定性的定义如下：

1°．初始值作微小变化，则是否解的变化也不大？

命 $(c_1,\cdots,c_n)$ 是 $t=t_0$ 时的初始值，因此初始值所得出的解是

$$x_i=g_i(t,c_1,\cdots,c_n).$$

命 $c_1^0,\cdots,c_n^0$ 是一组给定了的初始值．如果给了 $\varepsilon>0$，我们可以找出 $\delta\,(>0)$，使

$$|c_1-c_1^0|<\delta,\cdots,|c_n-c_1^0|<\delta$$

时，对所有的 $t\geqslant t_0$ 都有

$$|g_i(t,c_1,\cdots,c_n)-g_i(t,\ c_1^0,\cdots,\ c_n^0)|<\varepsilon$$

这样我们就说这个系统是**稳定的**．

并不失去普遍性，我们假定 $c_1^0=0,\cdots,c_n^0=0$，而且 $g_i(t,c_1^0,\cdots,c_n^0)=0$．这样我们的定义就变为特别简单了．当初始值

$$|c_i|<\delta,\quad i=1,2,\cdots,n$$

时，

$$|x_i(t)|<\epsilon,\quad t\geqslant t_0.$$

2°．加强条件：如果有一个 $\delta\,(>0)$ 存在当 $|c_i|<\delta$ 时，

$$\lim_{t\to\infty}x_i(t)=0$$

那末所研究的系统称为**渐近稳定的**．如果为了确切起见，我们加上“在 0 解附近”字样．

我们已知线性方程组

$$\frac{dx}{dt}=xA(t)$$

的解是

$$x=cX(t).$$

此处 c 是初始值，即当 $t=t_0$ 时，$x=c$，而 $X(t)$ 是适合于

$$\frac{dX}{dt}=XA(t),\quad X(t_0)=T$$

的方阵．

1) 如果 $X(t)$ 在 (t_0,∞) 中是有界方阵，即有 M 使

$$|x_{ij}(t)|\leqslant M,\quad (t\geqslant t_0,\ i,\ j=1,\cdots,\ n).$$

则

$$|x_i(t)|\leqslant Mn\max_{1\leqslant j\leqslant n}|c_j|.$$

因此取 $\delta<\dfrac{\varepsilon}{nM}$ 当 $|c_j|\leqslant\delta$ 时 $|x_i(t)|<\varepsilon$，也就是 0 解 $x_1=0,\cdots,x_n=0$ 所描写出的物理系统是稳定的．

2) $\lim_{t\to\infty} X(t) = 0$. 如此，则 $X(t)$ 当然有界，因而系统是稳定的. 再则由

$$\lim_{t\to\infty} x(t) = 0$$

可知系统是渐近稳定的.

3) 如果 $X(t)$ 无界. 即至少有一 $x_{ij}(t)$ 在 (t_0, ∞) 间没有上界，我们取 $c_1 = 0, \cdots, c_{i-1} = 0$, $c_i \neq 0$, $c_{i+1} = 0, \cdots, c_n = 0$ 如此则

$$x_j = c_i x_{ij}(t),$$

x_j 因而也无界，因此系统是不稳定的.

注意 系统可能是对某些初始值是稳定的而某些不稳。

§ 6. Ляпунов 变换

现在考虑一般的线性方程:

$$\frac{dx}{dt} = xA(t).$$

定义 1 变换

$$x = yL(t)$$

称为 Ляпунов 变换，如果方阵 $L(t)$ 适合于

1°. 在 (t_0, ∞) 中 $L(t)$ 有连续微商 $\frac{dL}{dt}$,

2°. 在 (t_0, ∞) 中 $L(t)$, $\frac{dL}{dt}$ 有界,

3°. 有一个常数 $m(>0)$ 存在，使 $L(t)$ 的行列式的绝对值 $> m$.

$L(t)$ 称为 Ляпунов 方阵.

例 1 常数满秩方阵是 Ляпунов 方阵.

例 2 如果 D 的特征数是纯虚数而且是单构的，则

$$L(t) = e^{Dt}$$

是 Ляпунов 方阵.

定理 1 Ляпунов 方阵之逆仍然是 Ляпунов 方阵.

定理 2 Ляпунов 变换使稳定性，渐近稳定性及非稳定性不变.

此二定理的证明不难.

再研究 y 所适合的微分方程: 由

$$yL(t)A = xA(t) = \frac{dx}{dt} = \frac{dy}{dt}L(t) + y\frac{dL}{dt}$$

可知

$$\frac{dy}{dt} = yL(t)A(t)L(t)^{-1} - y\frac{dL}{dt}L(t)^{-1}.$$

定义 1 如果有 Ляпунов 方阵 $L(t)$ 使

$$B(t) = L(t)A(t)L(t)^{-1} - \frac{dL}{dt}L^{-1}, \tag{1}$$

则 $A(t)$ 与 $B(t)$ 称为 Ляпунов 等价. 用符号

$$A \overset{\text{Л}}{=} B$$

记之.显然有

(i) $A \overset{\text{Л}}{=} A$.

(ii) 由 $A \overset{\text{Л}}{=} B$ 得 $B \overset{\text{Л}}{=} A$;由(1)得

$$A(t) = L(t)^{-1}B(t)L(t) + L^{-1}\frac{dL}{dt} = L^{-1}B(t)L - \frac{dL^{-1}}{dt}L.$$

(iii) 由 $A \overset{\text{Л}}{=} B$, $B \overset{\text{Л}}{=} C$ 可得 $A \overset{\text{Л}}{=} C$;由

$$C = MBM^{-1} - \frac{dM}{dt}M^{-1} = M\left(LAL^{-1} - \frac{dL}{dt}L^{-1}\right)M^{-1} - \frac{dM}{dt}M^{-1}$$

$$= MLA(ML)^{-1} - \left(M\frac{dL}{dt}L^{-1}M^{-1} + \frac{dM}{dt}M^{-1}\right)$$

$$= MLA(ML)^{-1} - \frac{d}{dt}(ML)(ML)^{-1}.$$

定义 2　凡与常系数方阵 Ляпунов 等价的方阵所对应的方程组,称为可化简组.

因此任一可化简组方程的解答一定是

$$X(t) = L(t)e^{At}$$

的形式,这儿 $L(t)$ 是 Ляпунов 方阵,A 是常数方阵. 至于解如上形的方程组一定是可化的那就不待证了.

§7. 周期性系数的微分方程组

定理 1　周期性系数的方程组一定是可化简组.

证明　假定

$$A(t+\tau) = A(t), \quad -\infty \leqslant t \leqslant \infty,$$

则

$$\frac{dX(t+\tau)}{dt} = X(t+\tau)A(t).$$

因此 $X(t+\tau)$ 也是一解,即得

$$X(t+\tau) = VX(t)$$

此处 V 是一个满秩的常数方阵.

由于 $|V| \neq 0$, 因此可以定义

$$V^{\frac{t}{\tau}} = e^{\frac{t}{\tau}\log V}$$

(注意 $\log V$ 所取的分支)取

$$L(t) = V^{-t/\tau}X(t)$$

这是一个周期函数,

$$L(t+\tau) = L(t),$$

而且 $|L(t)| \neq 0$, 因此 $L(t)$ 是一个 Ляпунов 方阵. 这样

我们定义

$$X(t) = Y(t)L(t),$$

$$Y(t) = V^{t/\tau}$$

因此

$$\frac{dY}{dt} = Y\left(\frac{1}{\tau}\log V\right).$$

再考虑

$$\frac{dx}{dt} = x(A(t) + B(t)), \tag{1}$$

$A(t)$ 是有周期性的，$\lim_{t\to\infty}\|B(t)\| = 0$ 的情况则可以用以上所述的 Ляпунов 变换变为一方程组其

$$A = |a_{ij}|_{1\leqslant i,j\leqslant n} = \frac{1}{\tau}\log V.$$

因而如 § 4 的方法处理.

如果方阵 V 的所有的特征根的绝对值都 <1，则方程组 (1) 的解是渐近稳定的. 如果这些根中有一个的绝对值 >1，则是非稳定的.

附记 1 因此

$$\frac{dX}{dt} = XA(t), \quad X(0) = I,$$

$(A(t+1) = A(t))$ 的解的形式是

$$X(t) = e^{Ct}Q(t),$$

此处 $Q(t+1) = Q(t)$，C 是常数方阵.

附记 2 有周期系数的微分方程的研究是一个重要而且艰难的课题. 例如，在数学物理中某些重要研究中所遇到 Mathieu 方程

$$\frac{d^2u}{dt^2} + (a + b\cos 2t)u = 0$$

及月球运动中所遇到一般的方程

$$\frac{d^2u}{dt^2} + \left[\sum_{n=0}^{\infty} a_n \cos nt + b_n \sin nt)\right] \qquad \text{(Hille)}$$

都是这类的. 有专著讨论.

§ 8. Ляпунов 等价

我们现在研究常数方阵的 Ляпунов 等价问题. 我们已经知道，如果

$$A \overset{\text{E}}{=} B, \text{ 即 } A = PBP^{-1},$$

则

$$A \overset{\text{Л}}{=} B.$$

因此任一方阵 A 一定 Ляпунов 等价于一 Jordan 标准型.

因此我们现在来考虑 Jordan 块的等价问题.

1) $n = 1$. Ляпунов 等价条件变为

$$b = a - \frac{dl}{dt}l^{-1} = a - \frac{d}{dt}\log l.$$

因此得出

$$l(t)=e^{(a-b)(t-t_0)}.$$

由于 $l(t)$ 及 $l(t)^{-1}$ 的有界性质可知

$$l(t)=ce^{i\lambda t},\ \lambda \text{ 是实数.}$$

即两个数 Ляпунов 等价，则他们的实数部分相等. 反过来，命 $a=\alpha+i\beta$，则取 $l(t)=e^{i\beta t}$，则

$$b=a-\frac{d}{dt}\log l=\alpha.$$

即一个数一定 Ляпунов 等价于它的实数部分.

2）一般的情况

$$aJ,\ a=\alpha+\beta i.$$

取 $L=e^{i\beta t}I$，就把它变为

$$\alpha J$$

的形状了.

因此得出

定理 1 （Еругин）. 每一个可化组都可以用 Ляпунов 变换 $X=LY$ 化为

$$\frac{dY}{dt}=YJ$$

此处 J 是特征数全为实数的 Jordan 方阵.

现在不去证明：如果不计 J 中对角线上诸子块的次序，这样的标准形是唯一的.

§9. 逼近于常系数的差分方程与微分方程

关于差分方程

$$x(t+1)=x(t)A(t),\ \lim_{t\to\infty}A(t)=A$$

参考：Фрейман, Успехи Мате. Надк, Том 12, 3, 243—264.

关于微分方程

$$\frac{dx}{dt}=xA(t),\ \lim_{t\to\infty}A(t)=A,$$

参考 R. Bellman，微分方程的解的稳定性理论 P.52—60.

第六章 二 次 型

§1. 凑 方

中学里学习二次方程时所用的“凑方”法有很多的用处，熟悉了这个方法可以处理不少问题，在讲初等微积分极大极小问题时已经讲过一些，在研究二次型的时候，这个方法更为重要.

n 个变数的二次型就是

$$Q=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j. \tag{1}$$

由于 $a_{ij}x_ix_j+a_{ji}x_jx_i=(a_{ij}+a_{ji})x_ix_j$，我们不妨假定

$$a_{ij}=a_{ji}.$$

对应于一个二次型，有一个方阵

$$A=(a_{ij}) \tag{2}$$

其中 $a_{ij}=a_{ji}$，即 $A=A'$，这是对称方阵，反之对应于一个对称方阵我们有一个二次型，二次型可以写成为

$$xAx', \tag{3}$$

此处 $x=(x_1,\cdots,x_n)$.

当 $x_1,\cdots,x_n$ 经过线性变换

$$x_i=\sum_{j=1}^{n}y_jp_{ji},\quad x=yP, \tag{4}$$

变为 $y_1,\cdots,y_n$ 时，二次型变为

$$\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}y_iy_j,\quad b_{ij}=b_{ji},\quad B=(b_{ij}),$$

其中 B 与 A 的关系如何？由 (1) 及 (3) 可知

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\sum_{s=1}^{n}y_sp_{si}\sum_{t=1}^{n}y_tp_{tj}$$
$$=\sum_{s=1}^{n}\sum_{t=1}^{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}p_{si}p_{tj}\right)y_sy_t,$$

即

$$b_{st}=\sum_{i=1}^{n}\sum_{j=1}^{n}p_{si}a_{ij}p_{tj}.$$

写成矩阵形式

$$B=PAP'. \tag{5}$$

定义 如果有一个满秩方阵 P 使(5)式成立，则 A 与 B 称为相合.

相合有等价关系.

我们的问题是；利用线性变换(4)，可以把二次型(1)化为怎样的形式。

假定 $a_{11} \neq 0$，把(1)写成为

$$Q = a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}y_1y_3 + \cdots + 2a_{1n}x_1x_n + a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n + \cdots + a_{nn}x_n^2.$$

凑方得

$$\begin{aligned} Q &= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 - a_{11}\left(\frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 \\ &\quad + a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n + \cdots + a_{nn}x_n^2 \\ &= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 + \frac{a_{22}a_{11} - a_{12}^2}{a_{11}}x_2^2 + 2\frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}}x_2x_3 \\ &\quad + \cdots + 2\frac{a_{11}a_{2n} - a_{12}a_{1n}}{a_{11}}x_2x_n + \frac{a_{33}a_{11} - a_{13}^2}{a_{11}}x_3^2 + \cdots \\ &\quad + 2\frac{a_{11}a_{3n} - a_{13}a_{1n}}{a_{11}}x_3x_n + \cdots + \frac{a_{nn}a_{11} - a_{1n}^2}{a_{11}}x_n^2. \end{aligned}$$

如果 $a_{22}a_{11} - a_{12}^2 \neq 0$，则还可以凑方，这样便不难得出

$$\begin{aligned} Q &= a_{11}(x_1 + \beta_{12}x_2 + \cdots + \beta_{1n}x_n)^2 + \frac{D\begin{pmatrix}1, & 2\\ 1, & 2\end{pmatrix}}{a_{11}}(x_2 + \beta_{23}x_3 + \cdots + \beta_{2n}x_n)^2 \\ &+ \frac{D\begin{pmatrix}1, & 2, & 3\\ 1, & 2, & 3\end{pmatrix}}{D\begin{pmatrix}1, & 2\\ 1, & 2\end{pmatrix}}(x_3 + \beta_{34}x_4 + \cdots + \beta_{3n}x_n)^2 + \cdots + \frac{D\begin{pmatrix}1, & \cdots, & n\\ 1, & \cdots, & n\end{pmatrix}}{D\begin{pmatrix}1, & 2, & \cdots, & n-1\\ 1, & 2, & \cdots, & n-1\end{pmatrix}}x_n^2. \qquad (6) \end{aligned}$$

此处

$$\beta_{ij} = \frac{D\begin{pmatrix}1, & 2, & \cdots, & i-1, & i\\ 1, & 2, & \cdots, & i-1, & j\end{pmatrix}}{D\begin{pmatrix}1, & 2, & \cdots, & i\\ 1, & 2, & \cdots, & i\end{pmatrix}}, \quad i < j.$$

命

$$\begin{aligned} y_1 &= x_1 + \beta_{12}x_2 + \beta_{13}x_3 + \cdots + \beta_{1n}x_n, \\ y_2 &= \qquad\quad x_2 + \beta_{23}x_3 + \cdots + \beta_{2n}x_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= \qquad\qquad\qquad\qquad\qquad x_n, \end{aligned}$$

则得以下的结果，如果 A 的主子行列式

$$D\begin{pmatrix}1\\ 1\end{pmatrix}, \quad D\begin{pmatrix}1, & 2\\ 1, & 2\end{pmatrix}, \cdots, D\begin{pmatrix}1, & \cdots, & n\\ 1, & \cdots, & n\end{pmatrix}$$

都不等于 0，则有一个三角方阵

$$P = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \beta_{12} & 1 & 0 & \cdots & 0 \\ \beta_{13} & \beta_{23} & 1 & \cdots & 0 \\ & \cdots & \cdots & \cdots & \end{pmatrix},$$

使

$$PAP'$$

变为对角线的形式，而且在求出 P 的过程中仅用加减乘除（不用开方及其他运算）。

如果二次型恒等于 0，则所有的系数都等于 0，何以见得？由

$$e_i A e_i' = a_{ii}, \qquad (e_i + e_j)A(e_i + e_j)' - e_i A e_i' - e_j A e_j' = 2a_{ij}$$

即可推出结论，这儿 e_i 是矢量$(0, \cdots, 0, 1, 0, \cdots, 0)$。第 i 个分量为 1，其他为 0。

因此对任一非恒等于 0 的二次型，必有一非 0 的系数。如果有一个 $a_{ii} \neq 0$，则可搬 x_i 换为 x_1 而如上法进行，如果 $a_{11} = \cdots = a_{nn} = 0$，而 $a_{ij} \neq 0$，不妨假定就是 $a_{12} \neq 0$。

变形 $x_1 = y_1$，$x_2 = y_1 + y_2$，$x_3 = y_3, \cdots, x_n = y_n$，就可以使

$$0x_1^2 + a_{12}x_1x_2 + 0x_2^2 + \cdots = a_{12}y_1^2 + \cdots + \cdots$$

中的 y_1^2 项的系数不等于 0，因此得

定理 1 任一对称方阵一定相合于一个对角线方阵，也就是任一二次型可以化为

$$\lambda_1 x_1^2 + \cdots + \lambda_r x_r^2 \quad \lambda_1, \cdots, \lambda_r \neq 0$$

的形式，而 r 是原方阵的秩。

在变化的过程中只用了四则运算。

在复数范围内，命

$$y_i = \sqrt{\lambda_i} x_i, \tag{7}$$

则得

$$Q = y_1^2 + \cdots + y_r^2. \tag{8}$$

即得

定理 2 在复数范围内，对称方阵一定相合于

$$\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

这儿 r 是原方阵的秩，因此任意两个等秩的对称方阵一定相合。

在实数范围内 $\sqrt{\lambda_i}$ 不一定是实的，因此不能化为(8)的形式，而必须分清那些 λ_i 是正的那些是负的，如果已知

$$\lambda_1 > 0, \cdots, \lambda_s > 0, \quad \lambda_{s+1} < , \cdots, \lambda_r < 0,$$

则由 $y_i = \sqrt{|\lambda_i|} x_i$ 可以把 Q 变为

$$Q = y_1^2 + \cdots + y_s^2 - y_{s+1}^2 - \cdots - y_r^2.$$

也就是在实数范围内 A 相合于

$$\begin{pmatrix} I^{(s)} & 0 & 0 \\ 0 & -I^{(r-s)} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9}$$

定义 正项项数减去负项项数（即 $s - (r - s) = 2s - r$）称为实二次型的标。

定理 3 在实数范围内，两个二次型相合的必要且充分条件是：他们的秩与标都相等，而(9)称为标准型。

前面已经证明，任一实二次型一定相合于以方阵(9)为系数的二次型，现在所待证明的是：如果

$$x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 = y_1^2 + \cdots + y_{s_1}^2 - y_{s_1+1}^2 - \cdots - y_{s_2}^2,$$

则 $s = s_1$，$r = r_1$，这儿 x，y 间有满秩的线性关系，秩应当相等，因此 $r = r_1$，现在仅需证明 $s = s_1$。

假定 $s < s_1$，考虑

$$x_1 = \cdots = x_s = 0, \quad y_{s_1+1} = \cdots = y_r = 0; \tag{10}$$

把 $x_1, \cdots, x_s$ 看成为 y 的线性型，则共有 $s + r_1 - s_1 < r$ 个方程式，共有 $x_{s+1}, \cdots, x_r, y_1, \cdots, y_{s_1}$ 个未知数，方程数 $s + r - s_1$ 小于未知数的个数，因此有一组非全为 0 的 $x_{s+1}, \cdots, x_r, y_1, \cdots, y_{s_1}$ 适合于 (10)，因此

$$-(x_{s+1}^2 + \cdots + x_r^2) = y_1^2 + \cdots + y_{s_1}^2$$

这是不可能的，因此 $s = s_1$.

§2. 大块凑方法

还是凑方的基本看法，把 A 划分为

$$A = \begin{pmatrix} A_1^{(s)} & L \\ L' & A_2^{(n-s)} \end{pmatrix}.$$

如果 $|A_1| \neq 0$，则

$$\begin{pmatrix} I & 0 \\ -L'A_1^{-1} & I \end{pmatrix} \begin{pmatrix} A_1 & L \\ L' & A_2 \end{pmatrix} \begin{pmatrix} I & 0 \\ -L'A_1^{-1} & I \end{pmatrix}' = \begin{pmatrix} A_1 & L \\ 0 & A_2 - L'A_1^{-1}L \end{pmatrix} \begin{pmatrix} I & -A_1^{-1}L \\ 0 & I \end{pmatrix}$$
$$= \begin{pmatrix} A_1 & 0 \\ 0 & A_2 - L'A_1^{-1}L \end{pmatrix}.$$

由大块凑方法，立刻看出

$$|A| = |A_1| \, |A_2 - L'A_1^{-1}L|.$$

定义 1　如果 $A' = -A$，则 A 称为斜对称，二斜对称方阵 A，B 称为相合，如果有满秩的 P，使

$$PAP' = B.$$

定理 1　A 一定相合于

$$\begin{pmatrix} F & 0 & \\ 0 & F & \\ & & \ddots \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

两斜对称方阵相合的必要充分条件是他们的秩相等.

证明　1）斜对称方阵的主对角线上的元素一定都等于0.

2）$n = 2$，由

$$\begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

所以定理当 $n = 2$ 时正确.

3）不妨假定 $a_{12} \neq 0$，把 A 划分为

$$\begin{pmatrix} A_1 & L \\ -L' & A_2 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}, \quad A_2 = A_2^{(n-2)}.$$

如此则

$$\begin{pmatrix} I & 0 \\ L'A_1^{-1} & I \end{pmatrix} \begin{pmatrix} A_1 & L \\ -L' & A_2 \end{pmatrix} \begin{pmatrix} I & 0 \\ L'A_1^{-1} & I \end{pmatrix}' = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 + L'A_1^{-1}L \end{pmatrix}.$$

而 $A_2 + L'A_1^{-1}L$ 是一 $n-2$ 行列的斜对称方阵，因而证明了本定理。

由此极易推得：

定理 2 奇次斜对称方阵一定是非满秩的。

定理 3 满秩的斜对称方阵之行列式是一个平方数。

附记 这儿我们仅用了四则运算，而没有用上开方。

§3. 仿射几何二次曲面的仿射分类

定义 命 $x=(x_1,\cdots,x_n)$, $y=(y_1,\cdots,y_n)$. 变形

$$x=yA+c,\quad |A|\neq 0 \tag{1}$$

称为仿射变换，此处 $A=A^{(n)}$, c 是一个 n 维矢量.

$x=(x_1,\cdots,x_n)$ 所成的空间称为仿射空间.

仿射几何学就是研究几何图形在仿射变换下的性质. 如果有两个图形，我有一个仿射变换把其一变为它一，则他们称为仿射等价.

仿射等价关系是一个等价关系.

例 1 任何一点一定仿射等价于另一点，这称为仿射空间的可递性，一直线变为一直线.

例 2 一个平面

$$xa'=\lambda$$

经仿射变换变为一个平面：

$$yAa'=\lambda-ca'.$$

任一平面可以变为

$$x_1=0.$$

例 3 任意不在同平面上 $n+1$ 个点可以变为

$$0,\ e_1,\ e_2,\cdots,\ e_n.$$

证明 先把一点变为 0. 其他各点 $x^{(i)}=(x_1^{(i)},\cdots,x_n^{(i)})(i=1,2,\cdots,n)$ 所成的方阵

$$A=\begin{pmatrix} x_1^{(1)}, & \cdots, & x_n^{(1)} \\ \cdots & \cdots & \cdots \\ x_1^{(n)}, & \cdots, & x_n^{(n)} \end{pmatrix}$$

是满秩的(由于 $n+1$ 点不在一平面上)，变换

$$x=yA$$

把 $x=x^{(i)}$ 变为 $y=e_i$.

例 4 同一直线上取三点 $x^{(1)}$, $x^{(2)}$, $x^{(3)}=tx^{(1)}+(1-t)x^{(2)}$, 他们经 (1) 变为

$$y^{(1)},\ y^{(2)},\ 及\ y^{(3)},$$

则

$$\begin{aligned} x^{(3)}&=tx^{(1)}+(1-t)x^{(2)}=t(y^{(1)}A+c)+(1-t)(y^{(2)}A+c) \\ &=(ty^{(1)}+(1-t)y^{(2)})A+c. \end{aligned}$$

即得

$$y^{(3)}=ty^{(1)}+(1-t)y^{(2)}.$$

由此推得：$x^{(1)}$, $x^{(3)}$ 间，$x^{(2)}$, $x^{(3)}$. 间距离之比等于 $y^{(1)}$, $y^{(3)}$ 间，$y^{(2)}$, $y^{(3)}$ 间距离之比。(注意经过仿射变换距离是可变的.)

二次曲面的一般形式是

$$xSx' + 2bx' + \gamma = 0. \tag{2}$$

此处 $S = S' = S^{(n)}$, b 是 n 矢量，γ 是常量，对应于一个二次曲面有一个 $n+1$ 行列的方阵

$$G = \begin{pmatrix} S & b' \\ b & \gamma \end{pmatrix}. \tag{3}$$

当 (2) 经过 (1) 而变为

$$\begin{aligned}&(yA + c)S(yA + c)' + 2b(yA + c)' + \gamma \\ &\quad = yASAy' + 2(cSA' + bA')y' + cSc' + 2bc' + \gamma.\end{aligned}$$

它所对应的方阵

$$\begin{aligned}F = \begin{pmatrix} T & a' \\ a & B \end{pmatrix} &= \begin{pmatrix} ASA' & A(Sc' + b') \\ (cS + b)A' & cSc' + 2bc' + \gamma \end{pmatrix} = \begin{pmatrix} A & 0 \\ c & 1 \end{pmatrix}\begin{pmatrix} S & b' \\ b & \gamma \end{pmatrix}\begin{pmatrix} A & 0 \\ c & 1 \end{pmatrix}' \\ &= \begin{pmatrix} A & 0 \\ c & 1 \end{pmatrix} G \begin{pmatrix} A & 0 \\ c & 1 \end{pmatrix}'.\end{aligned}$$

由此如果 G 与 F 所代表的二次曲面仿射等价，则 S 与 T 相合，G 与 F 相合．在复数范围内 S 与 T 的秩相等，G 与 F 的秩相等，在实数范围内 S 与 T 的标与秩相等，G 与 F 的秩与标也相等。

更具体些，在复数范围内，有 A 使

$$ASA' = \begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix},$$

即

$$x_1^2 + \cdots + x_r^2 + 2\sum_{i=1}^{n} b_i x_i + \gamma = 0.$$

凑方

$$(x_1 + b_1)^2 + \cdots + (x_r + b_r)^2 + 2\sum_{i=r+1}^{n} b_i x_i + \gamma' = 0, \qquad \gamma' = \gamma - \sum_{i=1}^{r} b_i^2.$$

分三种情况来研究

(i) 所有的 $b_i (i = r+1, \cdots, n)$ 都等于 0，而 $\gamma \neq 0$，则

$$x_1 + b_1 = \sqrt{\gamma'}\, y_1, \quad \cdots, \quad x_r + b_r = \sqrt{\gamma'}\, y_r$$

把 (2) 变为

$$y_1^2 + \cdots + y_r^2 + 1 = 0. \tag{4}$$

(ii) 所有的 $b_i (i = r+1, \cdots, n)$ 及 γ' 都等于 0，则 (2) 仿射等价于

$$y_1^2 + \cdots + y_r^2 = 0. \tag{5}$$

(iii) 并非所有的 $b_i (i = r+1, \cdots, n)$ 都等于 0，则由

$$x_1 + b_1 = y_1, \cdots, x_r + b_r = y_r, \quad \sum_{i=r+1}^{n} b_i x_i + \frac{1}{2}\gamma = y_{r+1}$$

$$y_1^2 + \cdots + y_r^2 + 2y_{r+1} = 0. \tag{6}$$

即在复数范围内，任何一个二次方曲面一定仿射等价于(4)，(5)，(6)之一，在这儿也同时证明了，G 的秩减 S 的秩 $\leqslant 2$.

定理 1 如果 S 的秩等于 G 的秩等于 r，则所对应的二次曲面仿射等价于 (5)；如果 S 的秩等于 G 的秩减1，则等价于(4)，如果 S 的秩等于 G 的秩减 2，则等价于(6).

再在实数范围，首先有 A 使

$$ASA' = \begin{pmatrix} I^{(r)} & 0 & 0 \\ 0 & -I^{(r-s)} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

即 (2) 实仿射等价于

$$x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 + 2\sum_{i=1}^{n} b_i x_i + \gamma = 0. \tag{7}$$

凑方，得

$$(x_1 + b_1)^2 + \cdots + (x_s + b_s)^2 - (x_{s+1} - b_{s+1})^2 - \cdots - (x_r - b_r)^2 + 2\sum_{i=r+1}^{n} b_i x_i + \gamma' = 0,$$

$$\gamma' = \gamma - \sum_{i=1}^{s} b_i^2 + \sum_{i=s+1}^{r} b_i^2.$$

(i) 如果 $b_i (i = r+1, \cdots, n)$ 及 r' 都等于 0，则(2)等价于

$$x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 = 0. \tag{8}$$

(ii) 如果 $b_i (i = r+1, \cdots, n)$ 都等于 0，而 $\gamma' \neq 0$，则由

$$x_1 + b_1 \to \sqrt{|\gamma'|}\, x_1, \cdots, \quad x_r - b_r \to \sqrt{|\gamma'|}\, x_r$$

可知 (2) 等价于

$$x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 = \pm 1. \tag{9}$$

(iii) 如果有一个 $b_i (i = r+1, \cdots, n) \neq 0$，则由

$$x_1 + b_1 \to x_1, \cdots, x_r - b_r \to x_r, \quad \sum_{i=r+1}^{n} b_i x_i + \frac{1}{2}\gamma' \to x_{r+1}$$

可知 (2) 等价于

$$x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 + 2x_{r+1} = 0. \tag{10}$$

由此得出结论

定理 2 如果 S 的秩等于 G 的秩，而 S 的标等于 $2s - r$，则二次曲面仿射等价于(8). 如果 G 的秩等于 S 的秩加一，则仿射等价于 (9). 如果 G 的秩等于 S 的秩加二，则仿射等价于(10).

特例：在实数平面上，即 $n = 2$ 时，二次曲线仿射等价于以下几种曲线之一

(i)	$x_1^2 + x_2^2 = 0$	点圆，
(ii)	$x_1^2 - x_2^2 = 0$	两条相交的直线，
(iii)	$x_1^2 = 0.$	一直条线，
(iv)	$x_1^2 + x_2^2 = 1$	圆，
(v)	$x_1^2 + x_2^2 = -1$	虚圆，
(vi)	$x_1^2 - x_2^2 = 1$	双曲线，
(vii)	$x_1^2 = 1$	两条平行的直线，
(viii)	$x_1^2 = -1$	两条虚直线，
(ix)	$x_1^2 = 2x_2$	抛物线.

在三维空间中，二次曲面仿射等价于（只考虑G满秩的情况）.

(i)	$x_1^2+x_2^2+x_3^2=0$,	点球，
(ii)	$x_1^2+x_2^2-x_3^2=0$,	圆锥，
(iii)	$x_1^2+x_2^2+x_3^2=1$,	球，
(iv)	$x_1^2+x_2^2+x_3^2=-1$,	虚球，
(v)	$x_1^2+x_2^2-x_3^2=1$,	单叶双曲面，
(vi)	$x_1^2-x_2^2-x_3^2=1$,	双叶双曲面，
(vii)	$x_1^2+x_2^2+x_3=0$,	椭圆抛物面，
(viii)	$x_1^2-x_2^2+x_3=0$,	双曲抛物面.

习题　试列出F非满秩的方程并且说明它们的几何名称.

§4. 射 影 几 何

定义　射影变换是以下形式的变换

$$y=(xA+b)/(xc+d). \tag{1}$$

此处 $A=A^{(n,n)}$, $b=b^{(1,n)}$, $c=c^{(n,1)}$, $d=d^{(1)}$. 并假定

$$P=\begin{vmatrix} A & c \\ b & d \end{vmatrix}\neq 0. \tag{2}$$

如果

$$z=\frac{(yA^*+b^*)}{(yc^*+d^*)},\quad P^*=\begin{pmatrix} A^* & c^* \\ b^* & d^* \end{pmatrix}$$

$$z=\frac{[(xA+b)A^*+b^*(xc+d)b^*]}{[(xA+b)c^*+(xc+d)d^*]}=\frac{(x(AA^*+cb^*)+bA^*+db^*)}{(x(Ac^*+cd^*)+bc^*+dd^*)}$$

也是一个射影变换，而且它的方阵

$$\begin{pmatrix} AA^*+cb^* & Ac^*+cd^* \\ bA^*+db^* & bc^*+dd^* \end{pmatrix}=\begin{pmatrix} A & c \\ b & d \end{pmatrix}\begin{pmatrix} A^* & c^* \\ b^* & d^* \end{pmatrix}.$$

即连续施行两次射影变换所得出的依然是射影变换，其对应方阵是原对应方阵的乘积，因此，对应于逆方阵，我们有变换(1)的逆变换.

注意　对应一个射影变换(1)，我们有一个方阵(2). 但是反过来，不同的方阵可以对应于同一的射影变换，例如 ρP 也是对应于(1)，这儿 ρ 是任一不等于 0 的数.

进一步，研究怎样的不同的P会对应于相同的变换(1)，我们只要考虑，怎样的P对应于恒等变换，在(1)中以

$$x=0,\qquad e_1,\cdots,e_n$$

代入，而要求

$$y=0,\qquad e_1,\cdots,e_n,$$

则得 $b=0$ 及

$$e_i(e_ic+d)=e_iA$$

e_ic+d 是一个数，因此得，$a_{ji}=0$，$(i\neq j)$，取 $x=y=\lambda e_i$，则对任 $\lambda\neq 0$，

$$(e_ic)\lambda^2e_i+\lambda de_i=\lambda a_{ii}e_i,\qquad (e_ic)\lambda+d=a_{ii}.$$

因此 $e_i c = 0$，即得 $c = 0$，并且 $a_{ii} = d$，即

$$P = [d, d, \cdots, d].$$

即当且仅当 $P = \rho I$ 时,(1)代表恒等变换.

不难证明,如果 P 与 Q 代表同一变换,则它们之间相差一个常数因子.

射影变换也是把线性关系变为线性关系.

仿射变换是射影变换,所以在射影变换下不变的性质,在仿射变换下也不变.

射影变换有以下的一些性质:

1）任意 $n+2$ 个点,其中任意 $n+1$ 点都不在一个 n 维的平面上,可以变为

$$0, e_1, \cdots, e_n, e_1 + \cdots + e_n.$$

证明　仿射变换已经把 $n+1$ 点变为

$$0, e_1 \cdots, e_n.$$

还有一点

$$(a_1, \cdots, a_n).$$

由假定 $a_1 \cdots a_n \neq 0$.（因为如果 $a_1 = 0$. 则此点在 $0, e_2, \cdots, e_n$ 所定的平面上). 射影变换

$$x_i = a_i y_i$$

把 $x_i = a_i$ 变为 $y_i = 1$. 即可得证.

射影变换的另一形式：命

$$\begin{pmatrix} A & c \\ b & d \end{pmatrix} \begin{pmatrix} A^* & c^* \\ b^* & d^* \end{pmatrix} = I,$$

则

$$(xc + d)^{-1}(xA + b) = (d^*x - b^*)(-c^*x + A^*)^{-1}.$$

要证明此式极易，因为它就等价于

$$\begin{aligned} &(xA + b)(-c^*x + A^*) - (xc + d)(d^*x - b^*) \\ &\quad = -x(Ac^* + cd^*)x + x(AA^* + cb^*) - (bc^* + dd^*)x + bA^* + db^* \\ &\quad = x - x = 0. \end{aligned}$$

假定 $x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}$ 经(1)而变为 $y^{(1)}, y^{(2)}, y^{(3)}, y^{(4)}$,

则

$$\begin{aligned} y^{(i)} - y^{(j)} &= (x^{(i)}c + d)^{-1}(x^{(i)}A + b) - (d^*x^{(j)} - b^*)(-c^*x^{(j)} + A^*)^{-1} \\ &= (x^{(i)}c + d)^{-1}[(x^{(i)}A + b)(-c^*x^{(j)} + A^*) - (x^{(i)}c + d)(d^*x^{(j)} - b^*)] \\ &\quad \times (-c^*x^{(j)} + A^*)^{-1} \\ &= (x^{(i)}c + d)^{-1}(x^{(i)} - x^{(j)})(-c^*x^{(j)} + A^*)^{-1}. \\ &= p^{(i)}(x^{(i)} - x^{(j)})Q^{(j)}. \end{aligned}$$

假定这些点在 x^*, x^{**} 的联线上，而且

$$x^{(i)} = t^{(i)}x^* + (1 - t^{(i)})x^{**}$$

并且假点 x^*, x^{**} 变为 y^*, y^{**}，而 $x^{(i)}$ 变为 $y^{(i)} = \lambda^{(i)}y^* + (1 - \lambda^{(i)})y^{**}$，则由

$$x^{(i)} - x^{(j)} = (t^{(i)} - t^{(j)})(x^* - x^{**}),$$

可知

$$(\lambda^{(i)}-\lambda^{(j)})(y^*-y^{**})=p^{(i)}(t^{(i)}-t^{(j)})(x^*-x^{**})Q^{(j)}.$$

x^i 与 x^j 的距离的平方等于 $(x^i-x^j)(x^i-x^j)'$，因此 x^i，x^j 之距离的平方与 y^i，y^j 之距离的平方的关系是:

$$(\lambda^{(i)}-\lambda^{(j)})^2(y^*-y^{**})(y^*-y^{**})'=(p^{(i)})^2(t^{(i)}-t^{(j)})^2$$
$$\cdot(x^*-x^{**})Q^{(j)}Q^{(j)\prime}(x^*-x^{**})'$$

$y^{(1)}$ 与 $y^{(2)}$ 的距离与 $y^{(3)}$ 与 $y^{(2)}$ 的距离的比等于

$$\cdot(\lambda^{(1)}-\lambda^{(2)})(\lambda^{(3)}-\lambda^{(2)})^{-1}=p^{(1)}(t^{(1)}-t^{(2)})(t^{(3)}-t^{(2)})^{-1}/p^{(3)}.$$

因此：得：

$$\frac{\lambda^{(1)}-\lambda^{(2)}}{\lambda^{(3)}-\lambda^{(2)}}\Big/\frac{\lambda^{(1)}-\lambda^{(4)}}{\lambda^{(3)}-\lambda^{(4)}}=\frac{t^{(1)}-t^{(2)}}{t^{(3)}-t^{(2)}}\Big/\frac{t^{(1)}-t^{(4)}}{t^{(3)}-t^{(4)}}.$$

这个

$$\frac{\lambda^{(1)}-\lambda^{(2)}}{\lambda^{(3)}-\lambda^{(2)}}\Big/\frac{\lambda^{(1)}-\lambda^{(4)}}{\lambda^{(3)}-\lambda^{(4)}}$$

称为四点的交比．由此可见射影变换使同直线四点的交比不变．

§5. 二次曲面的射影分类

二次曲面

$$xSx'+2bx'+\gamma=0,\quad F=\begin{pmatrix}S & b'\\ b & \gamma\end{pmatrix}$$

经射影变换变为

$$\begin{aligned}&(xA+b)S(xA+b)'+2(xc+d)b(xA+b)'+\gamma(xc+d)^2\\&=(xA+b)S(xA+b)'+(xc+d)b(xA+b)'\\&\quad+(xA+b)b'(xc+d)'+\gamma(xc+d)(xc+d)'\\&=x(ASA'+cbA'+Ab'c'+\gamma cc')x'+2(bSA'+dbA'\\&\quad+bb'c'+\gamma dc')x'+bSb'+2dbb'+\gamma d^2.\end{aligned}$$

它的方阵

$$G=\begin{pmatrix}ASA'+cbA'+Ab'c'+\gamma cc' & (bSA'+dbA'+bb'c'+\gamma dc')\\ bSA'+dbA'+bb'c'+\gamma dc' & bSb'+2dbb'+\gamma d^2\end{pmatrix}$$
$$=\begin{pmatrix}A & c\\ b & d\end{pmatrix}\begin{pmatrix}S & b'\\ b & \gamma\end{pmatrix}\begin{pmatrix}A & c\\ b & d\end{pmatrix}'=PFP'.$$

因此，二次曲面的射影分类特别容易．

在复数域中 F 的秩是唯一的不变量．任一秩等于 r 的 F 有一个 P 使

$$PFP'=\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix}.$$

即任一二次曲面一定射影等价于

$$x_1^2+\cdots+x_r^2=0,\qquad 如果\ r<n,$$
$$x_1^2+\cdots+x_r^2=1,\qquad 如果\ r=n.$$

前者称为退化的．因此非退化的二次曲面一定射影等价于球．

在实数域中，如果 F 是非退化的，即 F 是满秩的，那它一定等价于

$$x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_n^2=1,$$

在平面上非退化的二次曲线一定射影等价于实圆虚圆或双曲线。

§6. 定 正 型

现在实数范围内研究问题.

定义 1　对任一非 0 的矢量常有

$$xSx'>0, \quad (S'=S)$$

则 S 或其所对应的二次型称为定正的. 如果对所有的 x 常有

$$xSx'\geqslant 0$$

则 S 称为半定正的.

显然有定正的 S 一定是满秩的,因为如果 S 非满秩的,则有 $x(\neq 0)$ 使 $xS=0$, 因而 $xSx'=0$.

在实相合下,定正性(或半定正性)是不变的, 即如果 S 定正, 而 $T=PSP'$ (P 实, $|P|\neq 0$), 则 T 也是定正的.

由于任一对称方阵一定实相合于

$$\begin{pmatrix} I^{(s)} & 0 & 0 \\ 0 & -I^{(r-s)} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2).$$

因此, 半定正的条件是 $r=s$, 定正的条件是 $r=s=n$. 即如果 S 是半定正,则

$$S=P\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}P'.$$

在这表达式 P 的第 $r+1,\cdots,n$ 列都不起作用,如果把 P 的前 r 列写成为

$$Q=\begin{pmatrix} p_{11}, & \cdots, & p_{1r} \\ \cdots & \cdots & \cdots \\ p_{n1}, & \cdots, & p_{nr} \end{pmatrix}, \qquad Q=Q^{(n,r)},$$

则得

$$S=QQ'.$$

有两个有用的特例:

1) 如果 $r=1$, 则

$$S=u'u.$$

这儿 u 是一个 n 维实矢量.

2) 如果 $r=n$, 则任一定正的方阵可以表成为

$$S=PP'$$

的形式, 这儿 P 是满秩方阵.

反之也十分明显,因为 $y=xP$, 则

$$xSx'=yy'=\sum_{i=1}^{n}y_i^2=0.$$

即得 $y_1=\cdots=y_n=0$. 因而 $x_1=\cdots=x_n=0$.

定理 1 半定正方阵的主子方阵也是半定正的.

证明十分容易. 因为不妨假定主子方阵就是第 1，2…，l 行列的元素所成的方阵，即

$$S=\begin{pmatrix} S_1 & L \\ L', & S_2 \end{pmatrix}, \qquad S_1=S_1^{(l)}$$

中的 S_1・矢量 $x=(x_1,\cdots,x_l,0,\cdots,0)$ 使

$$xSx'=(x_1,\cdots,x_l)S_1(x_1,\cdots,x_l)'$$

由于 $xSx'\geqslant 0$，所以 S_1 是半正定的.

定理 2 半定正方阵的行列式 $\geqslant 0$.

这也是显然. 由于

$$\det(PP')=(\det P)^2\geqslant 0.$$

由此也可推出 Schwarz 不等式

$$\left|\begin{pmatrix} x \\ y \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}'\right|=\begin{vmatrix} xx' & xy' \\ yx' & yy' \end{vmatrix}=xx'yy'-(xy')^2\geqslant 0.$$

还有更广的不等式

$$\left|\begin{pmatrix} x \\ y \\ z \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}'\right|=\begin{vmatrix} xx' & xy' & xz' \\ yx' & yy' & yz' \\ zx' & zy' & zz' \end{vmatrix}\geqslant 0$$

等等.

§7. 用凑方法求最小值

1）假定 S 是定正的，C 是一个矢量，研究函数

$$F(x_1,\cdots,x_n)=xSx'+2xC'+\gamma \tag{1}$$

的最小值，

由

$$xSx'+2xC'+\gamma=(x+CS^{-1})S(x+CS^{-1})'+\gamma-CS^{-1}C'\geqslant\gamma-CS^{-1}C'$$

可知 $F(x_1,\cdots,x_n)$ 的最小值等于

$$\gamma-CS^{-1}C', \tag{2}$$

而且当且仅当 $x=-CS^{-1}$ 时，取此最小值.

如果 S 是定负的，则(1)有最大值. 如果 S 的标不等于 $\pm n$，则(1)无最大最小值. 因为 $S=P[1,1,\cdots,1,-1,\cdots,-1]P'$，命 $xP=y$，则

$$F(x_1,\cdots,x_n)=\sum_{i=1}^{s}y_i^2-\sum_{i=s+1}^{n}y_i^2+\cdots.$$

因之，当 $y_1\to\infty$ 时 $F\to\infty$，而 $y_n\to\infty$ 时 $F\to-\infty$（其他变数取常数可也).

2）条件极值.

在条件

$$xC'=\alpha$$

下求

$$xSx'$$

的最小值．考虑

$$xSx' + 2\lambda(xC' - \alpha)$$

的最小值，λ 是任一实数，由 1) 可知

$$xSx' + 2\lambda xC' - 2\lambda\alpha \geqslant -2\lambda\alpha - \lambda^2 CS^{-1}C' = \frac{\alpha^2}{CS^{-1}C'} - \left(\lambda + \frac{\alpha}{CS^{-1}C'}\right)^2 CS^{-1}C'.$$

当且仅当 $x = -\lambda CS^{-1}$ 时取等号．利用约束条件 $xC' = \alpha$，即 $-\lambda CS^{-1}C' = \alpha$ 可知 $xsx' \geqslant \frac{\alpha^2}{CS^{-1}C'}$，且仅当 $x = \frac{\alpha CS^{-1}}{CS^{-1}C'}$ 时取等号．

由此即得，对任一定正的 S 常有不等式

$$(xC')^2(=\alpha^2) \leqslant (xSx')(CS^{-1}C')$$

这又是 Schwarz 不等式的推广．因为如果 $S = I$ 这就是 Schwarz 不等式，读者试由 Schwarz 不等式来推出这不等式．

3）如果约束条件不止一个，即在条件

$$xC = \alpha$$

下求

$$xSx'$$

的最小值，此处 $C = C^{(n,l)}$，α 是一个 l 维的矢量．考虑函数

$$xSx' + 2(xC - \alpha)\lambda' = xSx' + 2xC\lambda' - 2\alpha\lambda'$$

这儿 λ 是一个 l 维的矢量．由 1) 可知

$$xSx' + 2(xC - \alpha)\lambda' \geqslant -2\alpha\lambda' - \lambda C'S^{-1}C\lambda'$$

且仅当 $x = -\lambda C'S^{-1}$ 时取等号．由约束条件

$$xC = -\lambda C'S^{-1}C = \alpha, \lambda = -\alpha(C'S^{-1}C)^{-1}$$

可得

$$xSx' \geqslant -\alpha\lambda' = \alpha(C'S^{-1}C)^{-1}\alpha'.$$

§ 8. Hessian

一个函数 $F(x_1, \cdots, x_n)$ 的 Hessian 是方阵

$$H(x_1, \cdots, x_n, F) = \left(\frac{\partial^2 F}{\partial x_i \partial x_j}\right)_{1 \leqslant i, j \leqslant n},$$

经变换

$$x_i = f_i(y_1, \cdots, y_n)$$

后

$$\frac{\partial F}{\partial x_i} = \sum_{k=1}^{n} \frac{\partial F}{\partial y_k} \frac{\partial y_k}{\partial x_i},$$

$$\frac{\partial^2 F}{\partial x_i \partial x_j} = \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{\partial^2 F}{\partial y_k \partial y_l} \frac{\partial y_k}{\partial x_i} \frac{\partial y_l}{\partial x_j} + \sum_{k=1}^{n} \frac{\partial F}{\partial y_k} \frac{\partial^2 y_k}{\partial x_i \partial x_j}.$$

即得

$$H(x_1, \cdots, x_n, F) = \left(\frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)}\right)' H(y_1, \cdots, y_n, F) \left(\frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)}\right)$$

$$+\sum_{k=1}^{n}\frac{\partial F}{\partial y_k}H(x_1,\cdots,x_n;\ y_k).$$

如果变换是线性的，

$$x_i=\sum_{i=1}^{n}y_j a_{ji},\qquad x=yA,$$

则

$$H(x_1,\cdots,x_n)=(A^{-1})'H(y_1,\cdots,y_n)A^{-1}.$$

定义．在区域D内，如果 $H(x_1,\cdots,x_n)$ 定正，函数F称为向下凹（成锅形），如果 $H(x_1,\cdots,x_n)$ 定负，则向上凸（成帽形）．

由上可知，线性变换使上凸或下凹的性质不变．

§9. 常系数二级偏微分方程分类

我们现在考虑常系数线性二级偏微分方程

$$\sum_{i,i=1}^{n}a_{ij}\frac{\partial^2 u}{\partial x_i\partial x_j}+2\sum_{i=1}^{n}b_i\frac{\partial u}{\partial x_i}+cu=0,\quad a_{ij}=a_{ji}\tag{1}$$

经变换

$$x_s=\sum_{t=1}^{n}y_t p_{ts},\quad x=yP\tag{2}$$

方程(1)变为

$$\sum_{s,t=1}^{n}a_{st}^*\frac{\partial^2 u}{\partial y_s\partial y_t}+2\sum_{s=1}^{n}b_s^*\frac{\partial u}{\partial y_s}+c^*u=0.\tag{3}$$

我们现在看其系数间的关系．

$$\frac{\partial u}{\partial y_t}=\sum_{i=1}^{n}\frac{\partial u}{\partial x_i}\frac{\partial x_i}{\partial y_t}=\sum_{i=1}^{n}\frac{\partial u}{\partial x_i}p_{ti}$$

及

$$\frac{\partial^2 u}{\partial y_s\partial y_t}=\sum_{i,j=1}^{n}\frac{\partial^2 u}{\partial x_i\partial x_j}p_{ti}p_{sj},$$

因此

$$\sum_{s,t=1}^{n}a_{st}^*\frac{\partial^2 u}{\partial y_s\partial y_t}+2\sum_{s=1}^{n}b_s^*\frac{\partial u}{\partial y_s}+c^*u$$
$$=\sum_{i,j=1}^{n}\frac{\partial^2 u}{\partial x_i\partial x_j}\sum_{s,t=1}^{n}p_{ti}a_{st}^*p_{sj}+2\sum_{i=1}^{n}\frac{\partial u}{\partial x_i}\sum_{s=1}^{n}b_s^*p_{sj}+c^*u$$

即得

$$a_{ij}=\sum_{s,t=1}^{n}p_{ti}a_{st}^*p_{si},$$

$$b_i=\sum_{s=1}^{n}b_s^*p_{si},$$

$$c=c^*,$$

即

$$A = P'A^*P, \quad b = b^*P,$$

而

$$A^* = P'^{-1}AP^{-1}, \quad b^* = bP^{-1}.$$

即 A^* 与 A 相合，所以在实数范围内，有 P 使

$$(P')^{-1}AP^{-1} = \begin{pmatrix} I^{(p)} & 0 & 0 \\ 0 & -I^{(r-p)} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

因此，(1)可以化为

$$\sum_{i=1}^{p} \frac{\partial^2 u}{\partial x_i^2} - \sum_{i=p+1}^{r} \frac{\partial^2 u}{\partial x_i^2} + 2\sum_{i=1}^{n} b_i \frac{\partial u}{\partial x_i} + cu = 0. \tag{4}$$

命

$$u = v e^{\alpha_1 x_1 + \cdots + \alpha_r x_r},$$

则

$$\frac{\partial u}{\partial x_i} = \left(\frac{\partial v}{\partial x_i} + \alpha_i v\right) e^{\alpha_1 x_1 + \cdots + \alpha_r x_r},$$

$$\frac{\partial^2 u}{\partial x_i^2} = \left(\frac{\partial^2 v}{\partial x_i^2} + 2\alpha_i \frac{\partial v}{\partial x_i} + \alpha_i^2 v\right) e^{\alpha_1 x_1 + \cdots + \alpha_r x_r},$$

代入(5)得

$$\sum_{i=1}^{p} \frac{\partial^2 v}{\partial x_i^2} - \sum_{i=p+1}^{r} \frac{\partial^2 v}{\partial x_i^2} + 2\sum_{i=1}^{p} (\alpha_i + b_i) \frac{\partial v}{\partial x_i} + 2\sum_{i=p+1}^{r} (-\alpha_i + b_i) \frac{\partial v}{\partial x_i}$$
$$+ 2\sum_{i=r+1}^{n} b_i \frac{\partial v}{\partial x_i} + c_0 v = 0,$$

可取 $\alpha_i = -b_i (i = 1, 2, \cdots, p)$，$\alpha_i = b_i (i = p+1, \cdots, n)$.

再取 P 使

$$(b_{r+1}, \cdots, b_n)P^{-1} = (1, 0, \cdots, 0).$$

因此得出标准型

$$\sum_{i=1}^{p} \frac{\partial^2 v}{\partial x_i^2} - \sum_{i=p+1}^{r} \frac{\partial^2 v}{\partial x_i^2} + \varepsilon \frac{\partial v}{\partial x_{r+1}} + cv = 0, \quad \varepsilon = 0 \text{ 或 } 1. \tag{5}$$

如果 $\varepsilon = 1$，可取 $v = we^{-cx_{r+1}}$ 因而可把 c 化为 0. 如果 $\varepsilon = 0$，可用 $x_i \to |c|^{-\frac{1}{2}} y_i$，而把 c 变为 ±1 因此我们有以下的几种标准型

(i) $$\sum_{i=1}^{p} \frac{\partial^2 u}{\partial x_i^2} - \sum_{i=p+1}^{r} \frac{\partial^2 v}{\partial x_i^2} = 0,$$

(ii) $$\sum_{i=1}^{p} \frac{\partial^2 v}{\partial x_i^2} - \sum_{i=p+1}^{r} \frac{\partial^2 v}{\partial x_i^2} \pm v = 0,$$

(iii) $$\sum_{i=1}^{p} \frac{\partial^2 v}{\partial x_i^2} - \sum_{i=p+1}^{r} \frac{\partial^2 v}{\partial x_i^2} + \frac{\partial v}{\partial x_{r+1}} = 0.$$

§ 10. Hermitian 型

现在考虑复数域，$\bar{a}$ 表示 a 的共轭复数。Hermitian 型是指以下的代数式。

$$\sum_{i,j=1}^{n} a_{ij}z_i\bar{z}_j,\quad a_{ij}=\bar{a}_{ji}.$$

所对应的方阵

$$H(a_{ij})$$

称为 Hermite 方阵．它适合于

$$H=\bar{H}'.$$

如果 $K=-\bar{H}'$ 则K称为斜 Hermite 方阵．显然可见 iK 是 Hermite 方阵．因此与对称方阵斜对称方阵的情况不同．斜 Hermite 方阵的研究并无什么新鲜之处的．

Hermitian 型可以写作

$$zH\bar{z}'.$$

经过变化

$$z=wP,$$

则得 Hermitian 型

$$wPH\bar{P}'\bar{w}$$

关系

$$H_1=PH\bar{P}^1$$

是一个等价关系，称为相联（Conjunctive）或称为H相合．用

$$H_1\underset{=}{H}H$$

表之．

显然如果A是 Hermite 方阵，而 $B\underset{=}{H}A$，则B也是，并有等价三性质：i）$A\underset{=}{H}A$，ii）由 $A\underset{=}{H}B$ 得 $B\underset{=}{H}A$，iii）由 $A\underset{=}{H}B$，$B\underset{=}{H}C$ 得 $A\underset{=}{H}C$．

与实域中对称方阵类似可以证明以下的定理．

定理 1 任意一个 Hermite 方阵一定H相合于

$$[1,1,\cdots,1-1,\cdots,-1,0,\cdots,0]$$

这儿有 p 个$+1$，q 个-1，$p+q=r$ 是H的秩，$p-q$ 也称为H的标．

由此推得

定理 2 两个 Hermite 方阵H相合的必要且充分条件是他们的秩相同，标相同．

定义 如果除 $z=0$ 以外常有

$$zH\bar{z}'>0$$

则H称为定正．如果常有

$$zH\bar{z}'\geqslant 0$$

则H称为半定正．

定正的必要且充分的条件是阶等于秩等于标．

§ 11. Hermitian 型的实形式

把z写成为 $x+iy$．把H写成为

$$S+iK,$$

这儿S是实对称方阵，K是实斜对称方阵．这样

$$zH\bar{z}' = (x+yi)(S+iK)(x'-iy') = (xS - yK + i(yS+xK))(x'-iy')$$
$$= (xS-yK)x' + (yS+xK)y' - [(xS-yK)y' - (yS+xK)x']i.$$

由于

$$yKy' = 0, \qquad xKx' = 0$$

及

$$xSy' = ySx'$$

虚数部分等于 0，因此

$$zH\bar{z}' = xSx' + 2xKy' + ySy' = (x, y)\begin{pmatrix} S & K \\ -K & S \end{pmatrix}(x,y)'.$$

因此 n 行列的 Hermite 方阵可以看成为 $2n$ 行列的特殊对称方阵

$$\begin{pmatrix} S & K \\ -K & S \end{pmatrix}.$$

而变形

$$z = wP$$

也可以写成为 $w = u + iv$, $P = A + iB$, 因此

$$x + iy = (u+iv)(A+iB) = uA - vB + i(vA + uB)$$

$$(x,y) = (u,v)\begin{pmatrix} A & B \\ -B & A \end{pmatrix}.$$

方阵

$$\begin{pmatrix} A & B \\ -A & B \end{pmatrix} \tag{1}$$

可以用以下的方法来刻划，任意一个与

$$\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{2}$$

可交换的方阵一定是形式 (1). 即由

$$\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

可得 $B = -C$, $A = D$.

因此 Hermite 方阵的理论可以说成为：在特殊变换下研究特殊对称方阵的性质的理论. 所谓特殊就是指形如 (1) 的方阵.

第七章　正交群与二次型对

§1. 正　交　群

我们现在考虑实矢量空间使矢量长度不变的线性变换，称为正交变换．即

$$y = x\Gamma, \tag{1}$$

使

$$xx' = yy'. \tag{2}$$

以 (1) 代入 (2) 得恒等式

$$x\Gamma\Gamma'x' = xx',$$

即得

$$\Gamma\Gamma' = I. \tag{3}$$

适合于 (3) 的方阵 Γ 称为正交方阵，显然连续运用两次正交变换仍得一正交变换，也就是两个正交方阵的乘积仍然是正交方阵．所有的 n 行列的正交变换(或方阵)称为正交群．由 (3) 可见 $|\Gamma| = \pm 1$.

我们现在在正交群下研究问题．正交变换使矢量的长度不变，也使二矢量的夹角余弦不变，即如果二矢量 $x^{(1)}$, $x^{(2)}$ 同时变为 $y^{(1)}$, $y^{(2)}$, 则

$$L(x^{(1)}, x^{(2)}) = \frac{x^{(1)}x^{(2)'}}{\sqrt{x^{(1)}x^{(1)'}x^{(2)}x^{(2)'}}}.$$

显然变为

$$\frac{y^{(1)}y^{(2)'}}{\sqrt{y^{(1)}y^{(1)'}y^{(2)}y^{(2)'}}}$$

正交条件 (3) 也可以改写成为：命 $\Gamma = (c_{ij})$, 则

$$\sum_{i=1}^{n} c_{ij}c_{ik} = \delta_{jk}. \tag{4}$$

这儿若 $j \neq k$, 则 $\delta_{jk} = 0$, 而 $\delta_{ii} = 1$, 由

$$\Gamma'\Gamma = I$$

可知 Γ' 也是正交的，因此还有一组与列有关的方程．由 (4) 也可以推得一个正交方阵是由 n 个两两正交的单位矢量所组成的：即

$$c^{(j)} = (c_{1j}, \cdots, c_{nj}).$$

再看 $n = 2$ 的情况，由

$$c_{11}^2 + c_{12}^2 = 1$$

可以取 $c_{11} = \cos\theta$, 如此则 $c_{12} = \pm\sin\theta$, 并不妨假定 $c_{12} = \sin\theta$ (若不然换 θ 为 $-\theta$ 即得)．由

$$c_{11}c_{21}+c_{12}c_{22}=0$$

可知 $c_{21}=-\rho\sin\theta$，$c_{22}=\rho\cos\theta$ 再由

$$c_{21}^2+c_{22}^2=1$$

可知 $\rho=\pm 1$。因此二维的正交方阵一定可以表成为

$$\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix},\ \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}.$$

前者是旋转，行列式=1，后者是反射，它的平方等于单位方阵，即

$$\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}^2=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

而行列式等于-1，反射是一个旋转乘以一个特殊反射

$$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

而成的.

一般的讲来，考虑由

$$a_{ii}=\cos\theta,$$
$$a_{ij}=\sin\theta,$$
$$a_{ji}=-\sin\theta,$$
$$a_{jj}=\cos\theta$$

及其他的

$$a_{st}=\delta_{st}$$

所定义的方阵是 (i,j) 平面上的旋转.

考虑
$$\begin{pmatrix}c_{11}, & \cdots, & c_{1n}\\ c_{21}, & \cdots, & c_{2n}\\ \cdots & \cdots & \cdots\\ c_{n1}, & \cdots, & c_{nn}\end{pmatrix}\begin{pmatrix}\cos\theta & \sin\theta & 0 & \cdots & 0\\ -\sin\theta & \cos\theta & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix}$$

可以取 θ 使 (2, 1) 地位的元素等于 0（即解 $c_{21}\cos\theta-c_{22}\sin\theta=0$ 即得）. 再乘以

$$\begin{pmatrix}\cos\theta & 0 & \sin\theta & 0 & \cdots & 0\\ 0 & 1 & 0 & 0 & \cdots & 0\\ -\sin\theta & 0 & \cos\theta & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & 0 & \cdots & 1\end{pmatrix},$$

可使 (3,1) 地位的元素等于 0，连续运用，可得一方阵其中 $c_{21}=\cdots=c_{n1}=0$，因此 $c_{11}=\pm 1$. 由于正交性质，$c_{12}=\cdots=c_{1n}=0$.

再乘以

$$\begin{pmatrix}1 & 0 & 0 & 0 & \cdots & 0\\ 0 & \cos\theta & \sin\theta & 0 & \cdots & 0\\ 0 & -\sin\theta & \cos\theta & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & 0 & & 1\end{pmatrix}$$

可以使 (3, 2) 地位的元素为 0 等. 因此，任一正交方阵可以乘以平面旋转方阵使它变为

$$[\pm 1, \pm 1, \cdots, \pm 1] \tag{5}$$

的形式.

又取 $\theta = \pi$

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

在 (5) 上乘以平面旋转的方阵可以得到

$$[1, 1, \cdots 1] \text{ 或 } [1, 1, \cdots, -1]$$

视其行列式是 +1 或 −1 而定.

总之，任何一个行列式为 1 的正交方阵可以表为平面旋转的乘积. 而行列式为 −1 的正交方阵可以表为以上的乘积外再乘以一个反射 $[1, 1, \cdots, -1]$.

定理 1 给了任意 m 个互相正交的单位矢量 $u^{(1)}, \cdots, u^{(m)}$, 我们一定可以找到 $n-m$ 个矢量 $u^{(m+1)}, \cdots, u^{(n)}$ 使

$$\begin{pmatrix} u^{(1)} \\ \vdots \\ u^{(n)} \end{pmatrix}$$

成一正交方阵.

证明 1) 先证,任一单位矢量 u 有正交方阵 Γ 使

$$u\Gamma = e_1.$$

这一点的证明极易. 因为可以定出 θ, 使

$$(u_1, u_2, \cdots, u_n) \begin{pmatrix} \cos\theta & \sin\theta & 0 & \cdots & 0 \\ -\sin\theta & \cos\theta & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

的第二个支量=0，续行此法可得所证.

2) 假定已知 $u^{(1)}\Gamma_1 = e_1$, 则

$$\begin{pmatrix} u^{(1)} \\ \vdots \\ u^{(m)} \end{pmatrix} \Gamma_1 = \begin{pmatrix} e_1 \\ v^{(2)} \\ \vdots \\ v^{(m)} \end{pmatrix}.$$

由于 $v^{(2)}, \cdots, v^{(m)}$ 与 e_1 正交,其第一支量皆等于 0, 再行此法可得

$$\begin{pmatrix} u^{(1)} \\ \vdots \\ u^{(m)} \end{pmatrix} \Gamma_1 \Gamma_2 \cdots \Gamma_m = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{pmatrix}$$

而

$$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \Gamma'_m \cdots \Gamma'_2 \Gamma'_1 = \begin{pmatrix} u^{(1)} \\ \vdots \\ u^{(n)} \end{pmatrix}$$

即得所求.

我们推广正交群之定义.

定义　命 S 是一定正对称方阵. 凡适合于

$$TST' = S$$

的方阵 T 称为 S 正交方阵. 显然 S 正交方阵乘积仍然是 S 正交方阵.

由于任何一个定正方阵可以表成为 $S = PP'$,因此

$$TPP'T' = PP',$$

即

$$(P^{-1}TP)(P^{-1}TP)' = I,$$

即 $P^{-1}TP$ 是正交方阵.

因此 S 正交方阵的研究并没有什么特别之处. 可由普通正交方阵的结果推得之.

但如果 S 非定正,则情况大有不同,其中如 S 仅有一个负号,就是所谓 Lorentz 群.这儿不谈.

§2. 定正二次型的平方根作为距离函数

在普通几何中，两点

$$x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n)$$

的距离等于

$$d(x,y) = \sqrt{(x-y)(x-y)'} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}.$$

这函数 $d(x,y)$ 适合于以下的四个条件:

i) 定正性，$d(x,y) \geqslant 0$，且仅当 $x = y$ 时取等号，

ii) 对称性：$d(x,y) = d(y,x)$，

iii) 等比性：$d(\lambda_x, \lambda_y) = |\lambda| d(x,y)$，

iv) 三角不等式

$$d(x,y) + d(y,z) \geqslant d(x,z) \tag{1}$$

而且仅当三点共线时取等号.

三角不等式的代数证明是：命 $x_i - y_i = a_i$，$y_i - z_i = b_i$，(1) 式等价于

$$\sqrt{aa'} + \sqrt{bb'} \geqslant \sqrt{(a+b)(a+b)'}.$$

平方之，它又等价于

$$aa'bb' \geqslant (ab')^2,$$

这就是 Буняковский-Schwarz 不等式.

余弦定律给与

$$d(x,z)^2 = d(x,y)^2 + d(y,z)^2 + 2d(x,y)d(y,z)\cos\theta,$$

即

$$\cos\theta = \frac{(x-y)(y-z)'}{\sqrt{(x-y)(x-y)'(y-z)(y-z)'}}, \tag{2}$$

这儿 θ 是 (x, y) 边与 (y, z) 边的夹角.

问题 是否还有其他的两点 x, y 的函数仍然有性质 i), ii), iii), iv). 这样的函数很多，最显然的例子是:

命 S 是一定正对称方阵，则

$$d(x,y)=\sqrt{(x-y)S(x-y)'}\ .$$

也适合于 i), ii), iii), iv)，这些性质的证明并不难. 只要写成 $S=PP'$，命 $\xi=xP$，$\eta=yP$，$\zeta=zP$，则立刻可由以上的 (1) 结果推出新性质来.

再举一个例子，当 $p\geqslant 1$ 时，

$$d(x,y)=\sqrt[p]{\sum_{i=1}^{n}|x_i-y_i|^p}\ .$$

§3. 空间的度量

在上节的研究中，不难看出 xSy' 起着十分重要的作用，这引出了以下的概念.

命 R 是实的 n 维矢量空间，我们定义其两个矢量 x, y 的内积(或称无向积). (x, y): 它是适合于以下性质的函数

1) $(x,y)=(y,x)$,

2) $(\lambda x,y)=\lambda(x,y)$, λ 是任意实数,

3) $(x+y,z)=(x,z)+(y,z)$.

由此立刻推出

2′) $(x,\lambda y)=\lambda(x,y)$,

3′) $(x,y+z)=(x,y)+(x,z)$, 函数 (x, x) 定义为 x 的范数，用 $N(x)$ 表之. 如果

4) $N(x)\geqslant 0$，而且仅当 $x=0$ 时，$N(x)=0$，这样的函数称为矢量空间的欧几里得度量.

有欧几里得度量的实矢量空间 R 称为欧几里得空间.

$\sqrt{(x,x)}$ 称为矢量的长度，长度等于 1 的矢量称为单位矢量或称为就范矢量. 任一矢量 x 可以化为单位矢量 $\dfrac{x}{\sqrt{(x,x)}}$.

如果 $(x,y)=0$，则称二矢量 x 与 y 互相正交，用 $x\perp y$ 表之.

如果 $(x,y)=0$，则由 1), 2), 3)，立刻推出

$$\begin{aligned}N(x+y,x+y)&=(x+y,x+y)=(x,x)+(y,y)\\&=N(x,x)+N(y,y).\end{aligned}$$

这就是商高定理.

这是 n 维的实矢量空间. 假定 $a^{(1)},\cdots\cdots,a^{(n)}$，是 n 个线性无关的矢量，则

$$x=\sum_{i=1}^{n}x_ia^{(i)},\quad y=\sum_{i=1}^{n}y_ia^{(i)}.$$

由 2), 3), 2′) 与 3′) 可知

$$(x,y)=\left(\sum_{i=1}^{n}x_ia^{(i)},\sum_{j=1}^{n}y_ja^{(j)}\right)=\sum_{i=1}^{n}x_i\left(a^{(i)},\sum_{j=1}^{n}y_ja^{(j)}\right)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j (a^{(i)}, a^{(j)}) = \sum_{i,j=1}^{n} S_{ij} x_i y_j,$$

此处 $(a^{(i)}, a^{(j)}) = S_{ij}$，而

$$N(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij} x_i x_j.$$

性质 4) 的意义为 $S = (S_{ij})$ 是定正的，因此这个抽象的处理所包含的具体内容，与定义

$$(x, y) = xSy'$$

无异. 虽然如此，这样的抽象定义却有很多的启发性。

定义 如果 $e_1, \cdots\cdots, e_m$，适合于

$$(e_i, e_j) = \delta_{ij} = \begin{cases} 0 \text{ 当 } i \neq j \\ 1 \text{ 当 } i = j, \end{cases} \quad i, j = 1, \cdots, m.$$

这 m 个矢量称为相互正交的单位矢量或称就范正交矢量，当 $m = n$ 时，称为就范正交基底.

由于 $S = PP'$，作

$$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = P^{-1} \begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(n)} \end{pmatrix},$$

则

$$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} S \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}' = I,$$

即 $(e_i, e_j) = \delta_{ij}$，因此 $e_1, \cdots, e_n$ 就是一组正交就范的基底.

§4. Gram-Schmidt 法

命 $x^{(1)}, \cdots, x^{(m)}$ 是 m 个任意线性无关的矢量. Gram-Schmidt 法是指从这 m 个矢量怎样求出这子空间的 m 个两两正交的单位矢量.

首先，取

$$u^{(1)} = \frac{1}{\sqrt{(x^{(1)}, x^{(1)})}} x^{(1)}$$

是一个单位矢量. 其次，作

$$u^{(2)} = \alpha u^{(1)} + \beta x^{(2)}.$$

如果它正交于 $u^{(1)}$，则得

$$0 = \alpha + \beta (x^{(2)}, u^{(1)}) = \alpha + \frac{\beta}{\sqrt{(x^{(1)}, x^{(1)})}} (x^{(2)}, x^{(1)}).$$

取

$$\alpha = -\beta \frac{(x^{(2)}, x^{(1)})}{\sqrt{(x^{(1)}, x^{(1)})}},$$

再则如果 $u^{(2)}$ 是单位矢量，则

$$\begin{aligned}1 = (u^{(2)}, u^{(2)}) &= (\alpha u^{(1)} + \beta x^{(2)}, \alpha u^{(1)} + \beta x^{(2)})\\ &= \alpha^2 + 2\alpha\beta(u^{(1)}, x^{(2)}) + \beta^2(x^{(2)}, x^{(2)})\\ &= \beta^2\left(-\frac{(x^{(1)}, x^{(2)})^2}{(x^{(1)}, x^{(1)})} + (x^{(2)}, x^{(2)})\right)\\ &= \beta^2\,\frac{(x^{(1)}, x^{(1)})(x^{(2)}, x^{(2)}) - (x^{(1)}, x^{(2)})^2}{(x^{(1)}, x^{(1)})},\end{aligned}$$

即得

$$\beta = \frac{\sqrt{(x^{(1)}, x^{(1)})}}{\sqrt{(x^{(1)}, x^{(1)})(x^{(2)}, x^{(2)}) - (x^{(1)}, x^{(2)})^2}},$$

$$\alpha = -\frac{(x^{(2)}, x^{(1)})}{\sqrt{(x^{(1)}, x^{(1)})(x^{(2)}, x^{(2)}) - (x^{(1)}, x^{(2)})^2}},$$

然后再取

$$u^{(3)} = \alpha u^{(1)} + \beta u^{(2)} + \gamma x^{(3)},$$

而从 $u^{(1)}u^{(3)\prime} = u^{(2)}u^{(3)\prime} = 0$，$u^{(3)}u^{(3)\prime} = 1$ 中定出 α, β, γ 来．继行之，最后得出

$$\begin{aligned}u^{(1)} &= \alpha_{11}x^{(1)},\\ u^{(2)} &= \alpha_{21}x^{(1)} + \alpha_{22}x^{(2)},\\ &\cdots\cdots\cdots\cdots\\ u^{(m)} &= \alpha_{m1}x^{(1)} + \alpha_{m2}x^{(m)} + \cdots\cdots + \alpha_{mm}x^{(m)}.\end{aligned}$$

这是 m 个两两正交的单位矢量．也就是

$$\begin{pmatrix}u^{(1)}\\ \vdots\\ u^{(m)}\end{pmatrix} = \begin{pmatrix}\alpha_{11} & 0 & \cdots & 0\\ \alpha_{21} & \alpha_{22} & & \\ \cdots & \cdots & \cdots & \cdots\\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mm}\end{pmatrix}\begin{pmatrix}x^{(1)}\\ \vdots\\ x^{(m)}\end{pmatrix}.$$

由

$$\begin{aligned}I^{(m)} &= \begin{pmatrix}u^{(1)}\\ \vdots\\ u^{(m)}\end{pmatrix} S \begin{pmatrix}u^{(1)}\\ \vdots\\ u^{(m)}\end{pmatrix}'\\ &= \begin{pmatrix}\alpha_{11} & 0 & \cdots & 0\\ \alpha_{21} & \alpha_{22} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mm}\end{pmatrix}\begin{pmatrix}x^{(1)}\\ \vdots\\ x^{(m)}\end{pmatrix} S \begin{pmatrix}x^{(1)}\\ \vdots\\ x^{(m)}\end{pmatrix}'\begin{pmatrix}\alpha_{11} & 0 & \cdots & 0\\ \alpha_{21} & \alpha_{22} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mm}\end{pmatrix}',\end{aligned}$$

可知以上的问题一变而为先作定正对称方阵

$$\begin{pmatrix}x^{(1)}\\ \vdots\\ x^{(m)}\end{pmatrix} S \begin{pmatrix}x^{(1)}\\ \vdots\\ x^{(m)}\end{pmatrix}'.$$

再用“凑方法”把它化为单位方阵；因此 α_{ij} 的数值在上一章中已经算出了！

这儿还须补充一点：

定理 1 (Gram)．M 个矢量 $x^{(1)}, \cdots, x^{(m)}$ 线性独立必要且充分条件是

$$\begin{pmatrix}x^{(1)}\\ \vdots\\ x^{(m)}\end{pmatrix} S \begin{pmatrix}x^{(1)}\\ \vdots\\ x^{(m)}\end{pmatrix}'$$

是定正的.

证明　首先这样形式的方阵总是半定正的. 原由是:

$$(u_1,\cdots,u_m)\begin{pmatrix}x^{(1)}\\ \vdots\\ x^{(m)}\end{pmatrix}S\begin{pmatrix}x^{(1)}\\ \vdots\\ x^{(m)}\end{pmatrix}'(u,\cdots u_m)'$$
$$=(u_1x^{(1)}+\cdots+u_mx^{(m)})S(u_1x^{(1)}\cdots+u_mx^{(m)})'.$$

如果这 m 个矢量是线性相关,则有非全为零的 $u_1,\cdots u_m$ 使 $u_1x^{(1)}+\cdots+u_mx^{(m)}=0$ 即方阵是半定正(而非定正的),反之,如果方阵是半定正的,则有非全为零的 $(u_1,\cdots,u_m)$ 使 $(u_1x^{(1)}+\cdots+u_mx^{(m)})S(u_1x^{(1)}+\cdots\cdots+u_mx^{(m)})'=0$。由 S 是定正推得

$$u_1x^{(1)}+\cdots+u_mx^{(m)}=0.$$

所以 $x^{(1)},\cdots,x^{(m)}$ 是线性相关的,定理证完.

§5. 正　投　影

假定 R 是欧几里得空间,假定 S 是一个子空间,这个子空间的基底是 $x^{(1)},\cdots,x^{(m)}$.我们要证明,任一矢量 x 可以表为

$$x=x_S+x_N,\quad x_S\in S,\quad x_N\perp S. \tag{1}$$

而 x_S 称为矢量 x 在子空间 S 上的正投影,而 x_N 称为射出矢量.

例如: R 是三维空间, S 是通过 0 的平面. 所有的矢量都是以原点为起点, x_S 是矢量在 S 平面上的正投影,而 x_N 是从矢量末端到平面 S 上的垂线. 而 $|x_N|$ 就是矢量末端到平面 S 的距离.

把 x_S 表成为

$$x_S=c_1x^{(1)}+\cdots+c_mx^{(m)},$$

此处 c 是实数.

从正交性出发

$$(x-x_S,x^{(k)})=0,\qquad (k=1,2,\cdots,m).$$

即

$$\begin{aligned}&(x^{(1)},x^{(1)})c_1+\cdots+(x^{(1)},x^{(m)})c_m+(x,x^{(1)})(-1)=0,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ &(x^{(m)},x^{(1)})c_1+\cdots+(x^{(m)},x^{(m)})c_m+(x,x^{(m)})(-1)=0,\\ &x^{(1)}c_1+\cdots+x^{(m)}c_m+x_S(-1)=0.\end{aligned}$$

由此推得

$$\begin{vmatrix}(x^{(1)},x^{(1)}),\cdots,(x^{(1)},x^{(m)}),x^{(1)}\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ (x^{(m)},x^{(1)}),\cdots,(x^{(m)},x^{(m)}),x^{(m)}\\ (x,x^{(1)}),\cdots,(x,x^{(m)}),x_S\end{vmatrix}=0,$$

即解得

$$x_S=-\frac{\begin{vmatrix}\Gamma & \begin{matrix}x^{(1)}\\ \vdots\\ x^{(m)}\end{matrix}\\ (x,x^{(1)}),\cdots,(x,x^{(m)}), & 0\end{vmatrix}}{\Gamma},$$

其中 $\Gamma(x^{(1)},\cdots,x^{(m)})$ 是矢量 $x^{(1)},\cdots,x^{(m)}$ 的 Gram 行列式. 而

$$x_N = x - x_S = \frac{\begin{vmatrix} & & & x^{(1)} \\ & \Gamma & & \vdots \\ & & & x^{(m)} \\ (x,x^{(1)}), & \cdots, & (x,x^{(m)}) & x \end{vmatrix}}{\Gamma}.$$

命 h 表矢量 x_N 的长度,则

$$h^2 = (x_N,x_N) = (x_N, x) = \frac{\begin{vmatrix} & & & (x^{(1)}, x) \\ & \Gamma & & \vdots \\ & & & (x^{(m)},x) \\ (x,x^{(1)}), & \cdots, & (x,x^{(m)}), & (x,x) \end{vmatrix}}{\Gamma} = \frac{\Gamma(x^{(1)},\cdots,x^{(m)}, x)}{\Gamma(x^{(1)},\cdots, x^{(m)})}.$$

这 h 有以下的几何意义:

$(m+1)$ 个矢量 $x^{(1)},\cdots,x^{(m)},x$ 作成一个 $(m+1)$ 维的单纯形,把由 $x^{(1)},\cdots,x^{(m)}$ 所成的 m 维单纯形作为底,由 x 的末端到这个底的高就等于 h.

假定 y 是 S 中的任一矢量,x 是 R 中的任一矢量,都是由原点引出的矢量. 由

$$N(x-y) = N(x_N + x_S - y) = N(x_N) + N(x_S - y) \geqslant N(x_N) = h^2,$$

可知

$$|x-y| \geqslant |x - x_S| = h,$$

即高不大于斜高,因此,在 S 中的所有的 y 中,用矢量 x_S 逼近的 x 的偏差最小. 而 $h = \sqrt{N(x-x_S)}$ 可以作为近似 $x \approx x_S$ 的测差平方.

引进 $\Gamma_p = \Gamma(x^{(1)},\cdots,x^{(p)}),\ p = 1, 2,\cdots, m$.

则

$$\sqrt{\Gamma_1} = |x^{(1)}| = V_1$$

是 $x^{(1)}$ 矢量的长度.

$$\sqrt{\Gamma_2} = V_1h_1 = V_2$$

是由 $x^{(1)}$, $x^{(2)}$ 所构成的平行四边形的面积,又

$$\sqrt{\Gamma_3} = V_2h_2 = V_3$$

是 $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ 为边所构成的平行六面体的体积. 继续定义

$$\sqrt{\Gamma_4} = V_3h_3 = V_4,$$

一般有

$$\sqrt{\Gamma_m} = V_{m-1}h_{m-1} = V_m.$$

很自然地,称 V_m 为由 $x^{(1)},\cdots,x^{(m)}$ 为边所构成的 m 维平行体的体积.

R 中取一正交就范基底,写

$$x^{(i)} = (x_{1i},\cdots,x_{ni}),\quad X = \begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(m)} \end{pmatrix},$$

则

$$\Gamma_m = |XX'|.$$

因而得

$$V_m^2 = \Gamma_m = \sum_{1\le i_1<i_2<\cdots<i_m\le n} \begin{vmatrix} x_{i_1 1}, & x_{i_1 2}, & \cdots, & x_{i_1 m} \\ x_{i_2 1}, & x_{i_2 2}, & \cdots, & x_{i_2 m} \\ \cdots & \cdots & \cdots & \cdots \\ x_{i_m 1}, & x_{i_m 2}, & \cdots, & x_{i_m, m} \end{vmatrix}^2 .$$

因此得出以下的几何性质:

定理 1 平行多面积的体积的平方等于在所有m维坐标空间上射影的体积的平方和.

特别是$m = n$,

$$V_n = \begin{vmatrix} x_{11}, & x_{12}, & \cdots, & x_{1n} \\ x_{21}, & x_{22}, & \cdots, & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1}, & x_{n2}, & \cdots, & x_{nn} \end{vmatrix} \text{ 的绝对值,}$$

再回到投影分解式

$$(x,x) = (x_S + x_N, x_S + x_N) = (x_S, x_S) + (x_N, x_N) \geqslant (x_N, x_N) = h^2.$$

即得

$$\frac{\Gamma(x^{(1)},\cdots,x^{(m)},x)}{\Gamma(x^{(1)},\cdots,x^{(m)})} = h_m^2 \leqslant (x,x) = \Gamma(x),$$

而且等号成立时的必要且充分的条件是 x 与 $x^{(1)},\cdots,x^{(m)}$ 正交. 由此立得 Hadamard 不等式

$$\Gamma(x^{(1)},\cdots,x^{(m)}) \leqslant \Gamma(x^{(1)})\Gamma(x^{(2)})\cdots\Gamma(x^{(m)}).$$

其中当且仅当 $x^{(1)},\cdots,x^{(m)}$ 两两正交时取等号,这等式的几何意义是:

平行多面体的体积不超过其边长的乘积,而且仅当其为长方形时等于这一乘积.

Hadamard 不等式的普通形式是:

$$\begin{vmatrix} x_{11}, & \cdots, & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{n1}, & \cdots, & x_{nn} \end{vmatrix}^2 \leqslant \sum_{i=1}^n x_{i1}^2 \sum_{i=1}^n x_{i2}^2 \cdots \sum_{i=1}^n x_{in}^2.$$

另证如次:

1) 如果 S, T 是定正,则 $S+T$ 也是定正,而 $|S+T| \geqslant |S|$. (化为 $S = I$, $T = [\lambda_1,\cdots,\lambda_n]$)

2) $$\begin{vmatrix} S_1 & L \\ L' & S_2 \end{vmatrix} = |S_1|\,|S_2 - LS_1^{-1}L'| \leqslant |S_1|\,|S_2|,$$

Hadamard 不等式乃其特例.

§6. 酉 空 间

在复矢量空间中,我们定义两个矢量 u, v 的内积为

$$u\overline{v'}.$$

使 $u\overline{u}'$ 不变的线性变换称为酉变换,即

$$v = uU,\quad U\overline{U}' = I.$$

而U也称为酉方阵.

当$n=1$时,酉方阵就是$e^{i\theta}$. 在n维空间

$$[e^{i\theta_1},\cdots,e^{i\theta_n}]$$

是酉方阵. 实正交方阵显然是酉方阵.

$U=(u_{ij})$ 有次之条件

$$\sum_{j=1}^{n} u_{ij}\bar{u}_{kj}=\delta_{ik}.$$

由 $\bar{U}'U=I$ 可知 $\bar{U}',U',\bar{U}$ 都是酉方阵.

命

$$u_{11}=\rho_{11}e^{i\theta_1},\cdots,u_{1n}=\rho_{1n}e^{i\theta_n}.$$

则

$$U[e^{-i\theta_1},\cdots,e^{-i\theta_n}]$$

的第一行是实数,即得一个方阵,其

$$\begin{pmatrix} u_{11}\cdots u_{1n}\\ \cdots\cdots \end{pmatrix},\quad u_{11}^2+\cdots+u_{1n}^2=1,$$

而且 $u_{11},\cdots,u_{1n}$ 是实的. 由上节已知有正交方阵Γ使

$$\begin{pmatrix} u_{11}\cdots u_{1n}\\ \cdots\cdots \end{pmatrix}\Gamma$$

的第一行是e_1, 依此续行可知: 酉方阵可以表为形如 $[e^{i\theta_1},\cdots,e^{i\theta_n}]$ 的酉方阵与正交方阵的乘积. 如果称 $[e^{i\theta_1},1,\cdots,1],\cdots[1,\cdots,e^{i\theta_i},1,\cdots,1]\cdots$为一维酉方阵, 则可知

定理 1 任一酉方阵可以表为平面旋转及一维酉方阵之积.

注意 1. $[-1,1]$是一个一维酉方阵.

2. 酉方阵的行列式是绝对值等于1的数. $|U\bar{U}'|=1,|U||\bar{U}|=1$, 故云.

定理 2 (Gram), m个复矢量

$$u^{(1)},\cdots,u^{(m)}$$

线性无关的必要且充分条件是 Hermite 方阵

$$(u^{(i)}\bar{u}^{(j)'})_{i,j=1,2,\cdots,m}$$

是定正的.

证法与前原则同,从略. 读者自己推广 Gram-Schmidt 方法.

同样不难证明, 如果有了m个矢量适合于

$$u^{(i)}\bar{u}^{(j)'}=\delta_{ij},\quad i,j=1,2,\cdots,m$$

则可以补充$n-m$个使

$$\begin{pmatrix} u^{(1)}\\ \vdots\\ u^{(n)} \end{pmatrix}$$

成为酉方阵.

把以上的结果作一些不难的推广.

在复矢量空间中定义内积 (u,v). 它适合于以下的一些性质

1) $(x,y)=\overline{(y,x)}$,

2) $(\alpha x,y)=\alpha(x,y)$,

3) $(x+y,z)=(x,z)+(y,z)$.

由此推得

2′) $(x,\alpha y)=\bar{\alpha}(x,y)$,

3′) $(x,y+z)=(x,y)+(x,z)$,

4) $Nx=(x, x)\geqslant 0$ 而且仅当 $x=0$ 时取等号．这样的空间称为酉空间．试推广以上所得的全部结果．

§7. 函数内积空间引

我们现在考虑在区间 (a, b) 上所有的可平方求积的实函数 $f(x)$ 类，称为 L^2 空间．我们定义

$$(f, g)=\int_a^b f(x)g(x)dx$$

为二函数 $f(x),g(x)$ 的内积，如果内积等于 0，此二函数称为正交．如果

$$\int_a^b f(x)^2dx=1,$$

则称为就范(或正常)．而

$$N(f)^{1/2}=\left(\int_a^b f^2(x)dx\right)^{1/2}$$

称为函数空间的度量．

如果有一组不全为零的实数 $c_1,\cdots, c_n$ 使

$$c_1f_1(x)+\cdots+c_nf_n(x)=0$$

几乎处处等于 0，则 $f_1(x),\cdots,f_n(x)$ 称为线性相关，不然称为线性独立．

定理 1 (Gram) n 个函数 $f_1(x),\cdots,f_n(x)$ 线性独立的必要且充分条件是对称方阵是

$$((f_i,f_j))_{1\leqslant i,j\leqslant n}$$

是定正的．

证明　考虑

$$(u_1,\cdots,u_n)\begin{pmatrix}\int_a^b f_1f_1dx, & \cdots, & \int_a^b f_1f_ndx\\ \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \int_a^b f_nf_1dx, & \cdots, & \int_a^b f_nf_ndx\end{pmatrix}(u_1,\cdots,u_n)'$$

$$=\left(\int_a^b(u_1f_1+\cdots+u_nf_n)f_1dx,\cdots\int_a^b(u_1f_1+\cdots+u_nf_n)f_ndx\right)\begin{pmatrix}u_1\\ \vdots\\ u_n\end{pmatrix}$$

$$=\int_a^b(u_1f_1+\cdots+u_nf_n)^2dx.$$

如果 $f_1,\cdots,f_n$ 是线性相关，则有非全为 0 的 $u_1,\cdots,u_n$ 使

$$\int_a^b(u_1f_1+\cdots+u_nf_n)^2dx=0.$$

反之，由此得出 $u_1f_1 + \cdots + u_nf_n$ 几乎处处为 0。

方阵

$$(f_i, f_j)_{1\leqslant i,j\leqslant n}$$

称为 gramian。

用凑方法有三角方阵使

$$\begin{pmatrix} \alpha_{11} & 0 & \cdots \\ \alpha_{21} & \alpha_{22} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix} (f_i, f_j)_{1\leqslant i,j\leqslant n} \begin{pmatrix} \alpha_{11} & 0 & \cdots \\ \alpha_{21} & \alpha_{22} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}' = 1,$$

如此得出一组函数

$$\begin{aligned} g_1 &= \alpha_{n}f_1, \\ g_2 &= \alpha_{21}f_1 + \alpha_{22}f_2, \\ &\cdots\cdots \\ g_m &= \alpha_{m1}f_2 + \alpha_{m2}f_1 + \cdots + \alpha_{mm}f_m, \\ &\cdots\cdots \end{aligned}$$

它们是正交而且就范的.

例 取 $a = -1$, $b = 1$。

函数 $1, x, x^2, x^3, \cdots$ 的 gramian 等于

$$\int_{-1}^{1} (1, x, x^2, x^3, \cdots)'(1, x, x^2, x^3, \cdots)dx$$

$$= 2\begin{pmatrix} 1 & 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \cdots \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & \cdots \\ \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix},$$

由于

$$1 = 1.1', \quad \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix}',$$

$$\begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{3} & 0 & \frac{2}{3\sqrt{5}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{3} & 0 & \frac{2}{3\sqrt{5}} \end{pmatrix}',$$

$$\begin{pmatrix} 1 & 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} \\ \frac{1}{3} & 0 & \frac{1}{5} & 0 \\ 0 & \frac{1}{5} & 0 & \frac{1}{7} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3\sqrt{5}} & 0 \\ 0 & \frac{\sqrt{3}}{5} & 0 & \frac{2}{5\sqrt{7}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3\sqrt{5}} & 0 \\ 0 & \frac{\sqrt{3}}{5} & 0 & \frac{2}{5\sqrt{7}} \end{pmatrix}',$$

因此得经 Gram-Schmidt 法所得出的正交就范函数是:

$$\frac{1}{\sqrt{2}}(1,x,x^2,\cdots)\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 & \cdots \\ \frac{1}{3} & 0 & \frac{2}{3\sqrt{5}} & 0 & \cdots \\ 0 & \frac{\sqrt{3}}{5} & 0 & \frac{2}{5\sqrt{7}} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}'^{-1},$$

以上的概念显然可作多种的推广:

1) 用 Lebesgue-Stieltje 积分的概念,命 $\alpha(x)$ 是 $[a, b]$ 中不递减的实函数,而且不是常数,则可以定义内积分

$$(f, g) = \int_a^b f(x)g(x)d\alpha(x).$$

2) 考虑复函数,而定义

$$(f, g) = \int_a^b f(x)\overline{g(x)}d\alpha(x).$$

3) 多维积分

$$(f, g) = \int\cdots\int f(x_1,\cdots,x_n)g(x_1,\cdots,x_n)d\alpha(x_1,\cdots,x_n).$$

4) 在 n 维空间的 m 维流形 S 上

$$(f, g) = \int\underset{\xi}{\cdots}\int f(x_1,\cdots,x_n)g(x_1,\cdots,x_n)d\sigma,$$

这儿 $d\sigma$ 是 ξ 上的测度.

5) 更抽象的所谓“内积空间”.

但一般说来,主要的方法的纲要就是从以上所讲的开始的.

以后我们还将择要介绍.

§8. 特 征 根

定理 1 任一方阵 A 可以唯一地表为

$$A = H + K, \quad H = \overline{H}', \quad K = -\overline{K}'.$$

命 a, h, k 代表 A, H, K 中元素的最大绝对值. 若 $\sigma = \alpha + \beta i$ 是 A 的特征根,则

$$|\sigma| \leqslant na, \quad |\alpha| \leqslant nh, \quad |\beta| \leqslant nk.$$

证明 由于 σ 是特征根,所以有一个非 0 矢量 z 使

$$zA = \sigma z.$$

由此可推出

$$\sigma z\bar{z}' = zA\bar{z}',$$

$$\bar{\sigma} z\bar{z}' = z\bar{A}'\bar{z}',$$

相加及相减得

$$\alpha z\bar{z}' = \frac{1}{2}z(A + \bar{A}')\bar{z}' = \sum_{i,j}\frac{1}{2}(a_{ij} + \bar{a}_{ji})z_i\bar{z}_j,$$

$$\beta z\bar{z}' = z\left(\frac{A-\bar{A}'}{2i}\right)\bar{z}' = \sum_{i,j}\frac{1}{2i}(a_{ij}-\bar{a}_{ji})z_i\bar{z}_j.$$

因此

$$|\sigma|z\bar{z}' \leqslant \sum_{i,j}|a_{ij}||z_i|\ |z_j| \leqslant a\sum|z_i||z_j| = a\left(\sum_i |z_i|\right)^2$$
$$\leqslant an\sum|z_i|^2 = anz\bar{z}'.$$

所以

$$|\sigma| \leqslant an.$$

同法

$$|\alpha|zz' \leqslant \sum_{i,j}|h_{ij}||z_i||\bar{z}_j| \leqslant h(\sum|z_i|)^2 \leqslant nhz\bar{z}'$$

及

$$|\beta|z\bar{z}' \leqslant nkz\bar{z}'.$$

这定理有以下的一些推理:

定理 2 Hermite 方阵的特征根是实的.

证明 $K=0$, 所以 $k=0$, 因此 $\beta=0$.

定理 3 实对称方阵的特征根是实的.

定理 4 斜对称方阵的特征根是纯虚的.

定理 5 如果定理 1 中的K是实的, 则

$$|\beta| \leqslant \sqrt{\frac{1}{2}n(n-1)}\,k.$$

证明 我们有

$$\beta z\bar{z}' = z\frac{A-\bar{A}'}{2i}\bar{z}' = \frac{1}{i}zK\bar{z}'$$

由于K是实的,所以K是实的斜对称方阵. 因此

$$\beta z\bar{z}' = \sum_{i<j}k_{ij}\frac{\bar{z}_iz_j - z_i\bar{z}_j}{i},$$

即

$$|\beta|z\bar{z}' \leqslant k\sum_{i<j}\left|\frac{\bar{z}_iz_j - z_i\bar{z}_j}{i}\right|.$$

再用 Schwarz 不等式

$$\left(\sum_{i<j}\left|\frac{\bar{z}_iz_j - z_i\bar{z}_j}{i}\right|\right)^2 \leqslant \frac{1}{2}n(n-1)\sum_{i<j}\left|\frac{\bar{z}_iz_j - z_i\bar{z}_j}{i}\right|^2$$
$$= -\frac{1}{2}n(n-1)\sum_{i<j}(\bar{z}_iz_j - z_i\bar{z}_j)^2$$
$$= \frac{n}{2}(n-1)((\sum z_i\bar{z}_i)^2 - \sum z_i^2\sum\bar{z}_i^2)$$
$$< \frac{n(n-1)}{2}(z\bar{z}')^2,$$

因此得

$$|\beta| \leqslant \sqrt{\frac{n(n-1)}{2}}k.$$

附记 这定理给出斜对称方阵特征根的较优上界.

定理 6 命 $\sigma = \alpha + i\beta$ 表示 $A = H + K$ 的特征根. 则有

$$m \leqslant \alpha \leqslant M,$$

这儿m是H的最小特征根,而M是最大特征根.

证明 命 $H = (h_{ij})$, 及 z 是对应于σ的A的特征矢量,则有

$$\alpha z\bar{z}' = zH\bar{z}',$$

即

$$\alpha = \frac{zH\bar{z}'}{z\bar{z}'}.$$

把 $zH\bar{z}'/z\bar{z}'$ 看为一个函数 $f(z)$. 则显然有

$$\min f(z) \leqslant \alpha \leqslant \max f(z).$$

现在只须证明 $\min f(z)$ 与 $\max f(z)$ 都是H的特征根即是,命 $z_j = x_j + iy_j$. 从

$$\frac{\partial f}{\partial x_j} = \frac{\partial f}{\partial y_j} = 0$$

立刻推得

$$\frac{\partial f}{\partial z_j} = 0, \qquad \frac{\partial f}{\partial \bar{z}_j} = 0$$

及

$$\frac{\partial f}{\partial z_j} = \frac{e_j H\bar{z}'}{z\bar{z}'} - \frac{e_j\bar{z}'}{(z\bar{z}')^2} zH\bar{z}' = 0,$$

极值 $\lambda(= zH\bar{z}'/z\bar{z}')$ 适合于

$$e_j(H - \lambda I)\bar{z}' = 0.$$

即

$$(H - \lambda I)\bar{z}' = 0.$$

即 $|H - \lambda I| = 0$.

由此也顺带地证明了.

定理 7 如果S是实对称,则最小(大)的特征根等于

$$\frac{xSx'}{xx'}$$

的最小(大)值. 也就是在球 $xx' = 1$ 上, xSx' 的最小(大值).

几何语言, $xSx' = \lambda$ 是一个椭球,而球 $xx' = 1$ 与椭球相切的情况是λ是S的特征根. 如果λ小于最小的特征根,则球全部在椭球之内. 如果λ大于最大的特征根,则球包有椭球.

定理 8 酉方阵的特征根的绝对值等于 1.

证明 如果x是U的特征根,则 $x + \frac{1}{x}$ 是Hermite 方阵 $U + U^{-1} = U + \bar{U}'$ 的特征根,而 $\frac{1}{i}\left(x - \frac{1}{x}\right)$ 是 $\frac{1}{i}(U - \bar{U}')$ 的特征根, 都是实的. 即

$$x + \frac{1}{x} = 2r,\ x - \frac{1}{x} = 2is\ (r, s \text{ 实数}),$$

则

$$x = r + is, \quad \frac{1}{x} = r - is.$$

因此 $r^2 + s^2 = 1$. 即得所证.

由此推得

定理 9 实正交方阵的复特征根成对出现.

§ 9. 积分方程的特征根

我们再把实对称的情况更简化地说一下：如果 α 是实对称方阵的特征根，则有一矢量 z 使

$$zS = \alpha z,$$

由此得

$$zS\bar{z}' = \alpha z\bar{z}'.$$

由于 $zS\bar{z}' = \overline{zS\bar{z}'}$ 是实数,所以 α 是实数.

用这个原则来处理积分方程的特征值问题.

假定 $K(x, y) = K(y, x)$ 是一个实连续函数. 其范围是 $a \leqslant x \leqslant b$, $a \leqslant y \leqslant b$. 求特征值 λ 使

$$f(x) = \lambda \int_a^b K(x,y)f(y)dy$$

有解 $f(x)$.

定理 1 特征值是实的.

证明 乘以 $\overline{f(x)}$ 而积分之得

$$\int_a^b |f(x)|^2 dx = \lambda \int_a^b \overline{f(x)}\, dx \int_a^b K(x,y)f(y)dy,$$

由于

$$\int_a^b \int_a^b \overline{K(x, y)\overline{f(x)}f(y)}\, dx\, dy = \int_a^b \int_a^b K(x,y)f(x)\overline{f(y)}\, dx\, dy$$
$$= \int_a^b \int_a^b K(y,x)\overline{f(x)}f(y)\, dx\, dy = \int_a^b \int_a^b K(x,y)\overline{f(x)}f(y)\, dx\, dy$$

是实的,因此 λ 也是实的.

习题 1 把本定理推广到复函数空间.

习题 2 试研究 $K(x,y) \neq K(y,x)$ 的情况,参考 § 5 之法.

§ 10. 对称方阵的正交分类

定理 1 在正交变换下，一个二次型可以变为

$$\lambda_1 x_1^2 + \cdots + \lambda_n x_n^2$$

的形式,也就是任何一个对称 S 有正交方阵 Γ 使

$$\Gamma S \Gamma' = [\lambda_1, \cdots, \lambda_n],$$

这儿 $\lambda_1, \cdots, \lambda_n$ 是实数,是 S 的特征根.

证明 对应于特征根 λ_1，有一特征矢量 x，不妨假定它就是单位矢量，即 $xx'=1$。

$$xS=\lambda_1 x.$$

有正交阵 Γ 化 x 为第一行，因此，$e_1\Gamma=x$，

$$e_1\Gamma S\Gamma'=\lambda_1 e_1,$$

所以 $\Gamma S\Gamma'$ 的形状是

$$\begin{pmatrix}\lambda_1 & 0\\ 0 & S_1\end{pmatrix},\quad S_1=S_1^{(n-1)}.$$

这证明了本定理。

由此定理可知

$$xSx'=\sum_{i=1}^{n}\lambda_i(\alpha_{i1}x_1+\cdots+\alpha_{in}x_n)^2,$$

此处 $\Gamma=(\alpha_{ij})$，即

$$S=\Gamma[\lambda_1,\cdots,\lambda_n]\Gamma',$$

而 $e_1\Gamma'$ 是 S 的特征矢量，即 $(\alpha_{11},\cdots,\alpha_{n1})$ 是 S 的特征根 λ_1 的特征矢量。

定理 2 对应于不相等特征根的特征矢量是正交的

证明 如果

$$x^{(1)}S=\lambda_1x^{(1)},\quad x^{(2)}S=\lambda_2x^{(2)},\quad \lambda_1\neq\lambda_2,$$

则

$$\lambda_1x^{(1)}x^{(2)\prime}=x^{(1)}Sx^{(2)\prime}=\lambda_2x^{(1)}x^{(2)\prime}.$$

由此推得 $x^{(1)}x^{(2)\prime}=0$。

由定理 1 不难立刻推得：把 λ_i 排成为

$$\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n,$$

则显然有

$$\frac{\lambda_1x_1^2+\cdots+\lambda_nx_n^2}{x_1^2+\cdots+x_n^2}\geqslant\lambda_1.$$

即

$$\min\frac{\lambda_1x_1^2+\cdots+\lambda_nx_n^2}{x_1^2+\cdots+x_n^2}=\lambda_1,$$

$$\min_{x_1=0}\frac{\lambda_1x_1^2+\cdots+\lambda_nx_n^2}{x_1^2+\cdots x_n^2}=\lambda_2.$$

换符号得：

定理 3 我们有

$$\min\frac{xSx'}{xx'}=\lambda_1,$$

当 $x=z^{(1)}$ 时取等号而

$$\min_{x\perp z^{(1)}}\frac{xSx'}{xx'}=\lambda_2,$$

且当 $x=z^{(2)}$ 时取等号；再

$$\min_{x\perp z^{(1)},z^{(2)}}\frac{xSx'}{xx'}=\lambda_3,$$

$$\cdots\cdots\cdots\cdots\cdots.$$

§ 11. 二次曲面的欧几里得分类

变换

$$y = x\Gamma + c$$

称为欧几里得变换，此处 Γ 是实正交方阵，c 是一实矢量．在旋转，反射之外，还有平移 $y = x + c$． 所有的欧几里得变换成一欧几里得群． 欧几里得几何学就是研究欧几里得群下的几何性质．不变性质．

首先是两点 $x^{(1)}, x^{(2)}$ 的距离是不变的；其理由是：

$$\begin{aligned}(y^{(1)} - y^{(2)})(y^{(1)} - y^{(2)})' &= (x^{(1)} - x^{(2)})\Gamma\Gamma'(x^{(1)} - x^{(2)})' \\ &= (x^{(1)} - x^{(2)})(x^{(1)} - x^{(2)})'.\end{aligned}$$

又任一超平面 $ax' = \lambda$ 经过欧几里得变换一定可以变为 $ax' = 0$．

现在看二次曲面

$$xAx' + 2xb' + \gamma = 0$$

的分类．

可以取得 Γ 使 $\Gamma A\Gamma' = (i_1, \cdots, i_n)$． 即经过 $x = y\Gamma$ 可变为

$$\sum_{i=1}^{n} \lambda_i y_i^2 + 2\sum_{i=1}^{n} b_i y_i + \gamma = 0,$$

如果诸 λ_i 皆不等于 0，则由

$$\sum_{i=1}^{n} \lambda_i \left(y_i + \frac{b_i}{\lambda_i}\right)^2 + \gamma - \sum_{i=1}^{n} \frac{b_i^2}{\lambda_i} = 0,$$

可知一般的二次曲面可以变为

$$\sum_{i=1}^{n} \lambda_i x_i^2 = \lambda.$$

如果 $\lambda \neq 0$，则可以变为

$$\sum_{i=1}^{p} \left(\frac{x_i}{a_i}\right)^2 - \sum_{i=p+1}^{n} \left(\frac{x_i}{a_i}\right)^2 = 1$$

的形式．

如果 $\lambda = 0$，即得退化二次曲面

$$\sum_{i=1}^{p} \left(\frac{x_i}{a_i}\right)^2 - \sum_{i=p+1}^{n} \left(\frac{x_i}{a_i}\right)^2 = 0.$$

如果有些 $\lambda_i = 0$，不妨取

$$\sum_{i=1}^{r} \lambda_i y_i^2 + 2\sum_{i=r+1}^{n} b_i y_i + \gamma = 0, \quad \lambda_1, \cdots, \lambda_r \neq 0.$$

如果非所有的 b_i 都等于0，我们可以把它再变为

$$\sum_{i=1}^{r} \lambda_i y_i^2 + y_{r+1} = 0.$$

总之，先看

$$\begin{pmatrix} A & b' \\ b & \gamma \end{pmatrix}.$$

如果是奇异的，则原二次曲面称为退化的，非退化的二次曲面一定欧几里得等价于以下的标准型

$$\sum_{i=1}^{p}\left(\frac{x_i}{a_i}\right)^2-\sum_{i=p+1}^{n}\left(\frac{x_i}{a_i}\right)^2=1. \tag{1}$$

$$\sum_{i=1}^{p}\left(\frac{x_i}{a_i}\right)^2-\sum_{i=p+1}^{n-1}\left(\frac{x_i}{a_i}\right)^2+x_n=0. \tag{2}$$

当 $p=n$ 时，所得出的曲面是椭球，而 $a_1,a_2,\cdots,a_n$ 是其主轴的长度。

读者试自己写出退化二次曲面的标准型。

§12. 方 阵 对

先复习一下方阵对的相抵关系。

方阵对 A_1, A_2 与 B_1, B_2 称为相抵，如果有二满秩方阵 P, Q 使

$$A_1=PB_1Q,\ A_2=PB_2Q$$

有如下的重要结果。

定理 1 如果 A_1, B_1 是满秩的，则方阵对相抵的必要且充分条件是 $A_1\lambda+A_2$ 与 $B_1\lambda+B_2$ 有相同的不变因子（或初等因子）。

此定理已经证过了。

我们现在研究 A_1, A_2 可能非满秩的情况。由较对称的形式

$$\lambda A_1+\mu A_2 \quad -\infty<\lambda,\mu<\infty$$

较为方便。如果有 λ,μ 使 $|\lambda A_1+\mu A_2|\neq 0$，则这些方阵称为非奇异方阵束。显然定理 1 可以推广到任意的非奇异束。

我们现在回来考虑对称方阵。

定理 2 如果

$$A=PBQ$$

及 A 与 B 都是对称的（或都是斜对称的）。P 与 Q 都是满秩的，则有一满秩方阵 R，与 P, Q 有关，但与 A, B 无关，使 $A=R'BR$.

证明 1) 由 $A=PBQ$，可知

$$A=Q'BP'=PBQ.$$

命 $U=Q'^{-1}P$，则

$$UB=BU'.$$

又 U 的任一多项式常有

$$f(U)B=B(f(U))'.$$

2) 满秩方阵 U 的平方根。我们现在证明有一多项式 $f(x)$ 存在，使

$$(f(U))^2=U.$$

如果这个对了，则命 $X=f(U)$，而

$$XB=BX',$$

取 $R=X'Q$，即得

$$R'BR=Q'XBX'Q=Q'X^2BQ=Q'UBQ=PBQ=A$$

即得所求。

命 $g(\lambda)$ 是U的特征多项式，如果能证明有 $f(\lambda)$ 存在使

$$f(\lambda)^2 \equiv \lambda(\mathrm{mod}\, g(\lambda))$$

也就证明了 2）所提出的结论。

3）将 $g(x)$ 分解为

$$g(x)=(x-a)^r(x-b)^s(x-c)^t\cdots$$

此处 $a, b, c, \cdots$两两不相等，且不为零。

把 $\sqrt{x}$ 展开成 $(x-a)$ 的幂级数

$$\sqrt{x}=\sqrt{a}+\frac{\sqrt{a}}{2a}(x-a)-\frac{(x-a)^2}{8a^2}+\cdots$$

命 $F(x)$ 为展开式的前 r 项之和．则$\sqrt{x}-F(x)$可被 $(x-a)^r$ 所整除．同样将 $\sqrt{x}$ 展开成 $(x-b)$ 的幂级数． $G(x)$ 为这一幂级数的前 s 项． 则 $\sqrt{x}-G(x)$ 可以被 $(x-b)^s$ 整除．如此等等。

把 $\dfrac{F(x)}{g(x)}$ 展开成部份分式得：

$$\frac{F(x)}{g(x)}=\frac{A_0}{(x-a)^r}+\cdots+\frac{A_{r-1}}{x-a}+R(x)=\frac{A(x)}{(x-a)^r}+R(x)$$

同样由 $G(x)$ 定义 $B(x)$ 等等，命

$$f(x)=g(x)\left(\frac{A(x)}{(x-a)^r}+\frac{B(x)}{(x-b)^s}+\cdots\right)$$

易知 $\sqrt{x}-f(x)$ 可以被 $g(x)$ 整除．即

$$x\equiv f(x)^2(\mathrm{mod}\, g(x))$$

由此极易推得

定理 3 如果

$$PAQ=A_1, \quad PBQ=B_1,$$

而A与A_1且都是对称的或斜对称的，B与B_1都是对称的或斜对称的，则有满秩的R使

$$R'AR=A_1, \quad R'BR=B_1.$$

定理 4 假定 $|A|\neq 0$，有满秩的P使

$$A=PBP'$$

的必要且充分条件是

$$\lambda A+\mu A' \text{ 与 } \lambda B+\mu B'$$

有相同的不变因子。

证明 1）若 $A=PBP'$，则

$$A'=PB'P',$$

所以

$$(\lambda A+\mu A')=P(\lambda B+\mu B')P'.$$

满足条件。

2）如果条件适合，则有

$$(\lambda A+\mu A')=P(\lambda B+\mu B')Q,$$

而

$$A+A'=P(B+B')Q,\quad A-A'=P(B-B')Q,$$

$A+A'$ 与 $B+B'$ 都是对称的，$A-A'$ 与 $B-B'$ 都是斜对称的，因此有 R 使

$$A+A'=R(B+B')R',A-A'=R(B-B')R'.$$

附记　当 $|A|=0$ 时，本定理仍正确，但证明较长不录.

§13. 斜对称方阵的正交分类

在实数范围内一对对称方阵的相合性，是并不简单的问题．我们已经讨论过的在正交群下二次型的分类就是这问题的一个特例．也就是方阵对

$$(I,S),(I,T)$$

的相合问题．可以稍加推广；如果方阵对中有一个是定正的，则可以化为

$$(I,[\lambda_1,\cdots,\lambda_n])$$

的标准形式．如果方阵束内有一个定正的，也可以同时化为对角线形式．这些推广是不难的．但如果没有定正的存在．问题变为很复杂．这儿不讨论了．

同样的情况对 Hermite 对也存在．有些书上错误地处理 Hermite 对的问题．

如果一个是对称 A_1，一个是斜对称 A_2，我们也只准备着重地讲一下 A_1 是定正的情况，也就不妨假定 $A_1=I$，即在正交群下，斜对称方阵 $A_2=K$ 的分类问题．

当 $n=2$ 时

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}'=(ad-bc)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

因此，任一二行二列的斜对称方阵一定正交等价于

$$\begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix},\quad k>0.$$

定理 1　任何一个斜对称方阵一定正交等价于

$$\begin{pmatrix} 0 & k_1 \\ -k_1 & 0 \end{pmatrix}\dotplus\begin{pmatrix} 0 & k_2 \\ -k_2 & 0 \end{pmatrix}\dotplus\cdots\dotplus\begin{pmatrix} 0 & k_\nu \\ -k_\nu & 0 \end{pmatrix}\dotplus 0\dotplus\cdots\dotplus 0$$

而 $k_1\geqslant k_2\geqslant\cdots\geqslant k_\nu>0$.

由前已知，斜对称方阵的特征根是纯虚的，即有一矢量 w 使

$$wK=ikw,$$

这儿 w 是复矢量．命 $w=u+iv$，则得

$$uK=-kv,\quad vK=ku$$

由 $uKu'=0$，可以推得 $uv'=0$，即 u 与 v 正交．

若 $k<0$，则由此推出

$$\begin{pmatrix} u \\ v \end{pmatrix}K=\begin{pmatrix} 0 & -k \\ k & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}$$

又

$$kuu'=vKu'=-uKv'=kvv',$$

所以 $uu'=vv'$ 命 $u_1=u/\sqrt{uu'}$，$v_1=v/\sqrt{vv'}$，则得

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} K = \begin{pmatrix} 0 & -k \\ k & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ v_1 \end{pmatrix},$$

而 $u_1u_1' = v_1v_1' = 1$，$u_1v_1' = 0$. 把 $\begin{pmatrix} u_1 \\ v_1 \end{pmatrix}$ 补充成为正交方阵 Γ，则得

$$\Gamma K \Gamma' = \begin{pmatrix} 0 & -k & 0 \\ k & 0 & \\ & 0 & K_1 \end{pmatrix}.$$

同样可以处理 $k > 0$ 的情况，依此续行，可以推出定理来●

§14. 辛群与辛分类

定义　使一个满秩斜对称方阵不变的方阵称为辛方阵，即如果 $K = -K'$，$|K| \neq 0$，使

$$PKP' = K$$

的方阵 P 称为辛方阵.

为了简单起见，我们可以固定

$$K = \begin{pmatrix} O & I \\ -I & O \end{pmatrix}, \quad I = I^{(V)}, \quad O = O^{(V)}.$$

把 P 写成为

$$P = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

则由

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} O & I \\ -I & O \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}' = \begin{pmatrix} O & I \\ -I & O \end{pmatrix}$$

可得

$$AB' = BA', \quad CD' = DC', \quad AD' - BC' = I.$$

立刻可以看出

$$\begin{pmatrix} I & S \\ O & I \end{pmatrix}, \begin{pmatrix} A & O \\ O & A'^{-1} \end{pmatrix}, \ S' = S, |A| \neq 0$$

都是辛方阵. 命 $J^2 = J$，而且是对角线方阵，即 J 的对角线上若干是 1 若干是 0. 则

$$\begin{pmatrix} J & I - J \\ -(I - J) & J \end{pmatrix}$$

也是辛方阵.

不难证明任一辛方阵也可以表为这些方阵之积.

§15. 各 式 分 类

我们有了一批群：正交群，辛群，酉群. 我们有了一批被分类的对象.

对称方阵，斜对称方阵，Hermite 方阵，正交方阵，辛方阵，酉方阵.

因此出现了一系列的问题，在某一个特定群下，把每种特定的方阵分类. 例如：对称方阵的辛分类，酉分类等等. 如果再加上"数的范围"就出现了种种的问题. 关于这些问题的专门研究，我不在此一一列举了.

但是我们必须指出，这样的问题中有不少是有实际意义的，而且是互相相关的.

例如：斜对称方阵的正交分类. 就可以用来处理正交方阵的正交分类，即求出正交方阵的标准型. 也就是给了一个正交方阵 Γ，我们一定有一个正交方阵 T 使

$$T\Gamma T^{-1}=\begin{pmatrix}\cos\theta_1 & \sin\theta_1\\ -\sin\theta_1 & \cos\theta_1\end{pmatrix}\dotplus\cdots\dotplus\begin{pmatrix}\cos\theta_\nu & \sin\theta_\nu\\ -\sin\theta_\nu & \cos\theta_\nu\end{pmatrix}\dotplus I.$$

证明 使

$$\Gamma=(I-K)(I+K)^{-1}.$$

由 $\Gamma\Gamma'=I$，可得 $K=-K'$.

由前已知有正交方阵 T 使

$$TKT'=\begin{pmatrix}0 & k_1\\ -k_1 & 0\end{pmatrix}\dotplus\cdots\dotplus\begin{pmatrix}0 & k_\nu\\ -k_\nu & 0\end{pmatrix}\dotplus 0\dotplus\cdots,$$

由此得出

$$T\Gamma T'=M_1^{(2)}\dotplus M_2^{(2)}\dotplus\cdots,$$

此处

$$M_i^{(2)}=\frac{I-\begin{pmatrix}0 & k\\ -k & 0\end{pmatrix}}{I+\begin{pmatrix}0 & k\\ -k & 0\end{pmatrix}}=\begin{pmatrix}1 & -k\\ k & 1\end{pmatrix}\begin{pmatrix}1 & k\\ -k & 1\end{pmatrix}^{-1}$$

$$=\begin{pmatrix}1 & -k\\ k & 1\end{pmatrix}^2\Big/(1+k^2)=\begin{pmatrix}\dfrac{1-k^2}{1+k^2} & -\dfrac{2k}{1+k^2}\\ \dfrac{2k}{1+k^2} & \dfrac{1-k^2}{1+k^2}\end{pmatrix}.$$

即得所求.

§16. 分 子 振 动

在质点系的经典力学处理时，动能等于

$$2T=\sum_{i,j=1}^{n}\dot{q}_i a_{ij}\dot{q}_j \quad a_{ij}=a_{ji},$$

位能等于

$$2V=\sum_{i,j=1}^{n}q_i b_{ij}q_j \quad b_{ij}=b_{ji},$$

这儿 q 是广义坐标，而 a,b 是与时间 t 及 q 无关的函数，这都是实数；$\dot{q}_i=\dfrac{dq_i}{dt}$.

命
$$A=(a_{ij}),\ B=(b_{ij})$$

及 $q=(q_1,\cdots,q_n)$，如此则得

$$2T=\dot{q}A\dot{q}' \tag{1}$$

$$2V=qBq'. \tag{2}$$

由于动能决不会是负的，而是仅当不动时才能为0，即仅当 $\dot{q}=0$ 时才能为0，因此 A 是一定正方阵.

我们已知有 P 使

$$PAP'=I, \tag{3}$$
$$PBP'=\Lambda.$$

再来考虑质点系的运动方程．这系的 Lagrangian 是

$$2L=2T-2V=\dot{q}A\dot{q}'-qBq',$$

因此运动方程是

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)-\frac{\partial L}{\partial q_i}=0,$$

写成方阵形式

$$\ddot{q}A+qB=0$$

命 $q=\theta P$，则得

$$\ddot{\theta}PAP'+\theta PBP'=0,$$

即

$$\ddot{\theta}+\theta\Lambda=0.$$

单开来写，就是

$$\frac{d^2\theta_i}{dt^2}+\lambda_i\theta_i=0,\quad i=1,2,\cdots,n.$$

如果是周期振动(或这系统的平衡位置是固定的)，这样 $\lambda_i>0$．即 B 也是定正的.

命 $\lambda_i=p_i^2$，如此得出

$$\theta_i=A_i\sin(p_it+\alpha_i).$$

再由 $q=\theta P$，我们得出

$$q_i=\sum_{k=1}^{n}A_k\sin(p_kt+\alpha_k)v_{ki}.$$

例　两个连在一起的调和振动.

命 x_1 代表质量 m_1 的物体从平衡位置向右移动的距离，而 x_2 是 m_2 移动后的距离，根据三弹簧的 Hooke 定理

$$2T=m_1\dot{x}_1^2+m_2\dot{x}_2^2,$$
$$2V=k_1'x_1^2+k_2'x_2^2+k_3(x_1-x_2)^2$$
$$=k_1x_1^2+k_2x_2^2-2k_3x_1x_2.$$

现在

$$A=\begin{pmatrix}m_1&0\\0&m_2\end{pmatrix},\quad B=\begin{pmatrix}k_1&-k_3\\-k_3&k_2\end{pmatrix}.$$

运动方程是

$$m_1\frac{d^2x_1}{dt^2}+k_1x_1-k_3x_2=0,\qquad m_1\frac{d^2x_2}{dt^2}-k_3x_1+k_2x_2=0.$$

现在

$$P=\begin{pmatrix}\frac{1}{\sqrt{m_1}}\cos\theta & \frac{1}{\sqrt{m_2}}\sin\theta\\ -\frac{1}{\sqrt{m_1}}\sin\theta & \frac{1}{\sqrt{m_2}}\cos\theta\end{pmatrix}.$$

这儿

$$\operatorname{tg}2\theta=\frac{2k_3/\sqrt{m_1m_2}}{k_2/m_2-k_1/m_1}.$$

因此

$$x_1=\frac{\theta_1}{\sqrt{m_1}}\cos\theta-\frac{\theta_2}{\sqrt{m_2}}\sin\theta,$$

$$x_2=\frac{\theta_1}{\sqrt{m_2}}\sin\theta+\frac{\theta_2}{\sqrt{m_1}}\cos\theta.$$

而

$$\frac{d^2\theta_1}{dt^2}+\lambda_1\theta_1=0,\quad \frac{d^2\theta_2}{dt^2}+\lambda_2\theta_2=0.$$

此处

$$\lambda_1=\omega_1^2=\frac{k_1}{m_1}\cos^2\theta-\frac{2k_3}{\sqrt{m_1m_2}}\cos\theta\sin\theta+\frac{k_2}{m_2}\sin^2\theta,$$

$$\lambda_2=\omega_2^2=\frac{k_1}{m_1}\sin^2\theta+\frac{2k_3}{\sqrt{m_1m_2}}\cos\theta\sin\theta+\frac{k_2}{m_2}\cos^2\theta.$$

第八章　体　　积

§1. m维流形的体积元素

在n维空间取m个以原点为始点的矢量

$$a^{(1)}, \cdots, a^{(m)}.$$

以这m个矢量为边作出m维的平行$2m$面体，前已算出，它的体积等于

$$|G|^{\frac{1}{2}}$$

这儿G是行列式（*Gramian*）

$$G = \left|\begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(m)} \end{pmatrix}\begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(m)} \end{pmatrix}'\right| \qquad (1)$$

的数值．

说得更清楚些，这个平行$2m$面体有以下的参变数的表达式：

$$x = \lambda_1 a^{(1)} + \cdots + \lambda_m a^{(m)}, \quad 0 \leqslant \lambda_\nu \leqslant 1. \qquad (2)$$

而G就是由距离

$$xx' = (\lambda_1, \cdots, \lambda_n)\begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(m)} \end{pmatrix}\begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(m)} \end{pmatrix}'(\lambda_1, \cdots \lambda_n)'$$

所导出的二次型的行列式的数值．

更一般地说，在n维空间有一m维的流形

$$x_i = x_i(\lambda_1, \cdots, \lambda_m), \quad i = 1, 2, \cdots, n. \qquad (3)$$

并假定其有连续微商，在一点有微分矢量

$$dx_i = \sum_{j=1}^{m} \frac{\partial x_i}{\partial \lambda_j} d\lambda_j,$$

即

$$dx = d\lambda \frac{\partial(x)}{\partial(\lambda)}.$$

所以在一点附近，这个m维流形的体积元素等于二次型

$$dxdx' = d\lambda\left(\frac{\partial(x)}{\partial(\lambda)}\right)\left(\frac{\partial(x)}{\partial(\lambda)}\right)' d\lambda'$$

的行列式的平方根，即

$$\left|\left(\frac{\partial(x)}{\partial(\lambda)}\right)\left(\frac{\partial(x)}{\partial(\lambda)}\right)'\right|^{\frac{1}{2}}.$$

用$d\sigma$表这m维流形的体积元素，则

$$d\sigma = \left|\left(\frac{\partial(x)}{\partial(\lambda)}\right)\left(\frac{\partial(x)}{\partial(\lambda)}\right)'\right|^{\frac{1}{2}} d\lambda_1 \cdots d\lambda_m.$$

因此，如果要求出 m 维流形 D^* 的体积，而当 $t\in D^*$ 时刻划出这块流形来，则体积等于

$$\int\cdots\int_{D^*}\left|\left(\frac{\partial(x)}{\partial(\lambda)}\right)\left(\frac{\partial(x)}{\partial(\lambda)}\right)'\right|^{\frac{1}{2}}d\lambda_1\cdots d\lambda_m. \tag{4}$$

换变数

$$\lambda_j=\lambda_j(t_1,\cdots,t_m),$$

则得

$$\frac{\partial(x)}{\partial(\lambda)}\ \frac{\partial(\lambda)}{\partial(t)}=\frac{\partial(x)}{\partial(t)}.$$

因此得出

$$\left|\frac{\partial(x)}{\partial(\lambda)}\left(\frac{\partial(x)}{\partial(\lambda)}\right)'\right|^{\frac{1}{2}}d\lambda_1\cdots d\lambda_m=\left|\frac{\partial(x)}{\partial(t)}\left(\frac{\partial(x)}{\partial(t)}\right)'\right|^{\frac{1}{2}}dt_1\cdots dt_m.$$

当然以上的叙述有不少需要严格化的地方，首先是 m 维流形的定义，我们必须引进所谓“区域坐标”. 因为，这个流形可能在某一范围内可以表为一种形式，而另一范围内另一种形式(不一定有统一的表示式). 其次必须定义“体积向”，如果以对 λ 的微分矢量作为基础，而改为变数 t 的时候，$\dfrac{\partial(\lambda)}{\partial(t)}$ 的正负号必须注意，如果是正号，体积是同向，不然是异向，也就是如果固定了：$e_1,\cdots,e_m$ 所成的平行 $2m$ 体的体积是正的，则

$$a_i^{(i)}=\sum_{j=1}^m p_{ij}e_j.$$

m 个矢量所成的平行 $2m$ 面体的体积. 当 $|p_{ij}|>0$ 时为正，而当 $|p_{ij}|<0$ 时为负.

我们还是先举些例子:

例 1　求 n 维球的表面积，即求

$$xx'=r^2$$

的表面积，微分之，使

$$xdx'=0$$

解出

$$x_1dx_1=-x_2dx_2-\cdots-x_ndx_n=-XdX'.$$

此处 $X=(x_2,\cdots,x_n)$. 因此

$$dxdx'=\left(\frac{XdX'}{x_1}\right)^2+dXdX'=dX\left(I+\frac{X'X}{x_1^2}\right)dX'.$$

这个二次型的行列式等于

$$\left|I+\frac{1}{x_1^2}X'X\right|=1+\frac{XX'}{x_1^2}=\frac{xx'}{x_1^2}=\frac{1}{x_1^2}.$$

(这儿用了 $|I+u'v|=1+uv'$.)　因此表面积等于

$$r\int\cdots\int_{r^2-XX'>0}\frac{dx_2\cdots dx_n}{|x_1|}=2^n r\int\cdots\int_{\substack{r^2-x_2^2\cdots-x_n^2>0\\ x_\nu\geqslant 0}}\frac{dx_2\cdots dx_n}{\sqrt{r^2-x_2^2\cdots-x_n^2}}.$$

这是一个 Dirichlet 积分，在下节中我们将算出它的数值等于

$$2\frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}r^{n-1}.$$

它等于 n 维球的体积

$$\frac{\pi^{\frac{n}{2}}}{\Gamma\left(1+\frac{n}{2}\right)}r^n$$

对半径的微商，读者思考一下，怎样的图形有此性质。

例 2　n 维空间的曲线

$$x_i = x_i(t),\quad a \leqslant t \leqslant b$$

的长度等于

$$\int_a^b \sqrt{dx\cdot dx'} = \int_a^b \sqrt{\sum_{i=1}^n \left(\frac{dx_i}{dt}\right)^2}\,dt.$$

例 3　球坐标系：

$$x = \rho u,\quad uu' = 1,$$
$$dx = d\rho u + \rho du'.$$

所以

$$dxdx' = d\rho^2 + 2\rho d\rho u du' + \rho^2 du du' = d\rho^2 + \rho^2 du du'.$$

球坐标：

$$u_1 = \cos\theta_1,$$
$$u_2 = \sin\theta_1\cos\theta_2,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$u_{n-1} = \sin\theta_1\cdots\sin\theta_{n-2}\cos\theta_{n-1},$$
$$u_n = \sin\theta_1\cdots\sin\theta_{n-2}\sin\theta_{n-1},$$
$$0 \leqslant \theta_i \leqslant \pi (i = 1,\cdots,n-2),\quad 0 \leqslant \theta_{n-1} \leqslant 2\pi.$$

因此

$$dxdx' = d\rho^2 + \rho^2(d\theta_1^2 + \sin^2\theta_1 d\theta_2^2 + \sin^2\theta_1\sin^2\theta_2 d\theta_3^2 + \cdots + \sin^2\theta_1\cdots\sin^2\theta_{n-2}d\theta_{n-1}^2).$$

所以球坐标的体积元素等于

$$\rho^{n-1}\sin^{n-2}\theta_1\sin^{n-3}\theta_2\cdots\sin\theta_{n-2}d\rho d\theta_1\cdots d\theta_{n-1}.$$

球表面的体积元素等于（ρ = 常数）

$$\rho^{n-1}\sin^{n-2}\theta_1\sin^{n-3}\theta_2\cdots\sin\theta_{n-2}d\theta_1\cdots d\theta_{n-1}.$$

（用此来回答例 1 所提出的问题）。

求球面上两点间的最短距离，不妨考虑单位球，而两点是

$$(1,0,\cdots,0),\ (\cos\alpha,\sin\alpha,0,\cdots,0)\quad 0 \leqslant \alpha \leqslant \pi.$$

因为经过旋转后任何球面上的两点一定可以达到这个地位，球面上任一曲线可以写成为 $\theta_\nu = \theta_\nu(t)$，$0 \leqslant t \leqslant 1$，由于过这二点，所以

$$\theta_1(0) = \theta_2(0) = \cdots = \theta_{n-1}(0) = 0,$$
$$\theta_1(1) = \alpha, \theta_2(1) = \cdots = \theta_{n-1}(1) = 0.$$

而这曲线的长度

$$\int_0^1 \sqrt{\left(\frac{d\theta_1}{dt}\right)^2 + \sin^2\theta_1\left(\frac{d\theta_2}{dt}\right)^2 + \cdots + \sin^2\theta_1\cdots\sin^2\theta_{n-2}\left(\frac{d\theta_{n-1}}{dt}\right)^2}\,dt \geqslant \int_0^1 \frac{d\theta_1}{dt}dt = \alpha.$$

而且仅当 $\theta_2 = 0$ 时取等号，即在 $(\cos\theta_1,\ \sin\theta_1,\ 0,\cdots,\ 0)$ 的平面上，因此，证明了，球面

上两点的距离以这两点与球心作平面交此球于"大弧"最短。

§ 2. Dirichlet 积 分

考虑 n 重积分

$$I=\int\cdots\int f(t_1+\cdots+t_n)t_1^{\alpha_1-1}\cdots t_n^{\alpha_n-1}dt_1\cdots dt_n.$$

求积分的范围是

$$t_\nu\geqslant 0,\quad t_1+\cdots+t_n\leqslant 1.$$

如果 f 是连续函数，而且 $\alpha_\nu>0(\nu=1,2,\cdots,n)$，则积分 I 可以化为单积分。

先简化积分

$$F(\lambda)=\int_0^{1-\lambda}\int_0^{1-\lambda-T}f(t+T+\lambda)t^{\alpha-1}T^{\beta-1}dtdT,$$

命 $t=T(1-v)/v$，则

$$F(\lambda)=\int_0^{1-\lambda}\int_{T/(1-\lambda)}^{1}f\left(\lambda+\frac{T}{v}\right)(1-v)^{\alpha-1}v^{-\alpha-1}T^{\alpha+\beta-1}dvdT$$
$$=\int_0^1\int_0^{(1-\lambda)v}f\left(\lambda+\frac{T}{v}\right)(1-v)^{\alpha-1}v^{-\alpha-1}T^{\alpha+\beta-1}dTdv$$

(换积分次序).

再命 $T=v\tau_2$，则得

$$F(\lambda)=\int_0^1\int_0^{1-\lambda}f(\lambda+\tau_2)(1-v)^{\alpha-1}\tau_2^{\alpha+\beta-1}v^{-\alpha-1+(\alpha+\beta-1)+1}d\tau_2dv$$
$$=\int_0^{1-\lambda}f(\lambda+\tau_2)\tau_2^{\alpha+\beta-1}d\tau_2\int_0^1(1-v)^{\alpha-1}v^{\beta-1}dv$$
$$=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\int_0^{1-\lambda}f(\lambda+\tau_2)\tau_2^{\alpha+\beta-1}d\tau_2.$$

因此，

$$I=\underset{\substack{t_3+\cdots+t_n\leqslant 1\\ t_\nu\geqslant 0}}{\int\cdots\int}t_3^{\alpha_3-1}\cdots t_n^{\alpha_n-1}dt_3\cdots dt_n$$
$$\times\int_0^{1-t_3-\cdots-t_n}\int_0^{1-t_2-\cdots-t_n}f(t_1+\cdots+t_n)t_1^{\alpha_1-1}t_2^{\alpha_2-1}dt_1dt_2$$
$$=\underset{\substack{t_3+\cdots+t_n\leqslant 1\\ t_\nu\geqslant 0}}{\int\cdots\int}t_3^{\alpha_3-1}\cdots t_n^{\alpha_n-1}dt_3\cdots dt_n$$
$$\times\frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}\int_0^{1-t_3-\cdots-t_n}f(t_3+\cdots+t_n+\tau_2)\tau_2^{\alpha_1+\alpha_2-1}d\tau_2$$
$$=\frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}\underset{\substack{\tau_2+t_3+\cdots+t_n\leqslant 1\\ t_\nu\geqslant 0\,\tau_2\geqslant 0}}{\int\cdots\int}f(\tau_2+t_3+\cdots+t_n)\tau_2^{\alpha_1+\alpha_2-1}t_3^{\alpha_3-1}\cdots t_n^{\alpha_n-1}d\tau_2dt_3\cdots dt_n$$

仿此续行，最后得

$$I=\frac{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_n)}{\Gamma(\alpha_1+\cdots+\alpha_n)}\int_0^1 f(\tau)\tau^{\alpha_1+\cdots+\alpha_n-1}d\tau,$$

这就是 Dirichlet 积分。

例 1　求单纯形

$$x = \lambda_1 a^{(1)} + \cdots + \lambda_m a^{(m)},$$
$$\lambda_1 + \cdots + \lambda_m \leqslant 1, \quad \lambda_v \geqslant 0$$

的体积.

$$dxdx' = d\lambda \begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(m)} \end{pmatrix} \begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(m)} \end{pmatrix}' d\lambda'.$$

因此

$$\int\cdots\int_{\substack{\lambda_1+\cdots+\lambda_m\leqslant 1 \\ \lambda_v\geqslant 0}} \left| \begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(m)} \end{pmatrix} \begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(m)} \end{pmatrix}' \right|^{\frac{1}{2}} d\lambda_1\cdots d\lambda_m = |G|^{\frac{1}{2}} \frac{\Gamma(1)^m}{\Gamma(m)} \int_0^1 \tau^{m-1} a\tau$$
$$= |G|^{\frac{1}{2}} \frac{1}{\Gamma(m+1)} = \frac{1}{m!} |G|^{\frac{1}{2}}.$$

即得 m 维单纯形的体积等于张此形的 m 个矢量的 ganmian 的平方根，再除以 $m!$

另证提示：

$$1 = \int_0^1 \cdots \int_0^1 dx_1 \cdots dx_m = m! \int\cdots\int_{0\leqslant x_1\leqslant\cdots\leqslant x_m\leqslant 1} dx_1\cdots dx_m = m! \int\cdots\int_{\substack{t_1+\cdots+t_m\leqslant 1 \\ t_v\geqslant 0}} dt_1\cdots dt_m.$$

$$(x_1 = t_1, x_2 = x_1 + t_2, \quad x_3 = x_2 + t_3, \cdots)$$

例 2

$$J = \int\cdots\int_{\sum_{i=1}^{n}|x_i|^p\leqslant r^p} dx_1\cdots dx_n = \frac{(2r)^n \Gamma\left(1+\frac{1}{p}\right)^n}{\Gamma\left(1+\frac{n}{p}\right)}.$$

把此积分分为 2^n 块，其每一块的积分等于

$$\int\cdots\int_{\substack{\sum_{v=1}^{n} x_v^p\leqslant r^p \\ x_v\geqslant 0}} dx_1\cdots dx_n = r^n \int\cdots\int_{\substack{\sum_{v=1}^{n} x_v^p\leqslant 1 \\ x_v\geqslant 0}} dx_1\cdots dx_n,$$

换变数 $x_v^p = y_v$，则

$$\frac{1}{2^n} J = r^n \int\cdots\int_{\substack{\sum_{v=1}^{n} y_v\leqslant 1 \\ y_v\geqslant 0}} \left(\frac{1}{p}\right)^n y_1^{\frac{1}{p}-1} \cdots y_n^{\frac{1}{p}-1} dy_1\cdots dy_n$$
$$= \frac{r^n \Gamma\left(\frac{1}{p}\right)^n}{p^n \Gamma\left(\frac{n}{p}\right)} \int_0^1 \tau^{\frac{n}{p}-1} d\tau = \frac{r^n \Gamma\left(1+\frac{1}{p}\right)^n}{\Gamma\left(1+\frac{n}{p}\right)},$$

特别取 $p=2$，则 r 为半径的 n 维球的体积等于

$$\frac{\pi^{\frac{n}{2}} r^n}{\Gamma\left(1+\frac{n}{2}\right)} = \vartheta_n r^n.$$

例 3

$$\int\cdots\int_{xx'\leqslant r^2} x'x dx_1\cdots dx_n = \frac{\vartheta_n}{n+2} r^{n+2} I.$$

如果 $i \neq j$，易证

$$\int\cdots\int_{xx'\leqslant r^2} x_i x_j dx_1\cdots dx_n = 0$$

(即换 $x_i \to -x_i$ 即得)，另一方面

$$\int\cdots\int_{xx'\leqslant r^2} x_1^2 dx_1\cdots dx_n = r^{n+2}2^{-n} \int\cdots\int_{\substack{y_1+\cdots+y_n\leqslant 1 \\ y_v\geqslant 0}} y_1^{\frac{1}{2}} y_2^{-\frac{1}{2}}\cdots y_n^{-\frac{1}{2}} dy_1\cdots dy_n$$

$$= r^{n+2} \frac{\Gamma\left(\frac{1}{2}\right)^{n-1} \Gamma\left(\frac{3}{2}\right)}{2^n \Gamma\left(1+\frac{n}{2}\right)} \int_0^1 \tau^{\frac{n}{2}} d\tau = r^{n+2} \frac{\pi^{\frac{n}{2}}}{\Gamma\left(1+\frac{n}{2}\right)} \frac{1}{n+2}.$$

即得所求.

§3. 正态分布积分

定理 1 如果 S 是定正的，则

$$\int_{-\infty}^{\infty}\cdots\int e^{itx'-\frac{1}{2}xSx'} dx_1\cdots dx_n = \frac{(2\pi)^{\frac{n}{2}}}{\sqrt{|S|}} e^{-\frac{1}{2}tS^{-1}t'}.$$

证明 当 $S=I$ 时，

$$\int_{-\infty}^{\infty}\cdots\int e^{itx'-\frac{1}{2}xx'} dx_1\cdots dx_n = \prod_{v=1}^{n} \int_{-\infty}^{\infty} e^{it_v x_v - \frac{1}{2}x_v^2} dx_v$$

$$= \prod_{v=1}^{n} \left(\sqrt{2\pi} e^{-\frac{1}{2}t_v^2}\right) = (2\pi)^{\frac{n}{2}} e^{-\frac{1}{2}tt'}.$$

在一般情况下，命

$$S = TT', \quad y = xT; \quad t = uT'^{-1},$$

则得

$$\int_{-\infty}^{\infty}\cdots\int e^{itx'-\frac{1}{2}xSx'} dx_1\cdots dx_n = \frac{1}{|T|}\int_{-\infty}^{\infty}\cdots\int e^{iuy'-\frac{1}{2}yy'} dy_1\cdots dy_n$$

$$= \frac{(2\pi)^{\frac{n}{2}}}{|S|^{\frac{1}{2}}} e^{-\frac{1}{2}uu'} = \frac{(2\pi)^{\frac{n}{2}}}{\sqrt{|S|}} e^{-\frac{1}{2}tS^{-1}t'}.$$

定理 2 S 仍然假定是定正的，u 与 v 是两个线性无关的矢量，则

$$\int_{\substack{-\infty \\ ux'\geqslant 0 \\ vx'\geqslant 0}}^{\infty}\cdots\int e^{-\frac{1}{2}xSx'} dx_1\cdots dx_n = \frac{(2\pi)^{\frac{n}{2}-1}}{\sqrt{|S|}} \cos^{-1}\frac{-uS^{-1}v'}{\sqrt{uS^{-1}u'vSv'}}.$$

证明 把 S 写成为 $S=TT'$，$xT\to x$，$uT'^{-1}\to u$，$vT'^{-1}\to v$，则原等式一变而为

$$\int_{\substack{-\infty \\ ux'\geqslant 0 \\ vx'\geqslant 0}}^{\infty}\cdots\int e^{-\frac{1}{2}xx'}dx_1\cdots dx_n=(2\pi)^{\frac{n}{2}-1}u\hat{\ }v, \tag{1}$$

这儿 $u\hat{\ }v$ 表示二矢量 u, v 的夹角.

显然可见,不妨假定 u, v 都是单位矢量，有正交方阵 Γ 使

$$u\Gamma=e_1,\quad v\Gamma=(\cos\alpha,\ \sin\alpha,\ 0,\cdots,0).$$

这儿 $\alpha=u\hat{\ }v$. 命 $x=y\Gamma'$, 则得

$$\int_{\substack{-\infty \\ ux'\geqslant 0 \\ vx'\geqslant 0}}^{\infty}\cdots\int e^{-\frac{1}{2}xx'}dx_1\cdots dx_n=\int_{\substack{-\infty \\ y_1\geqslant 0 \\ \cos\alpha y_1+\sin\alpha y_2\geqslant 0}}^{\infty}\cdots\int e^{-\frac{1}{2}yy'}dy_1\cdots dy_n$$

$$=\int_{\substack{-\infty \\ y_1\geqslant 0 \\ \cos\alpha y_1+\sin\alpha y_2\geqslant 0}}^{\infty}\cdots\int e^{-\frac{1}{2}(y_1^2+y_2^2)}dy_1dy_2(2\pi)^{\frac{1}{2}(n-2)}.$$

命 $y_1=\rho\sin\theta$, $y_2=\rho\cos\theta$, 则

$$\int_{\substack{-\infty \\ y_1\geqslant 0 \\ -\cos\alpha y_1+\sin\alpha y_2\geqslant 0}}^{\infty}\int e^{-\frac{1}{2}(y_1^2+y_2^2)}dy_1dy_2=\int_0^{\infty}e^{-\frac{1}{2}\rho^2}\rho d\rho\int_0^{\alpha}d\theta=\alpha.$$

即得所求，即得 (1).

(1) 式的几何说明是函数 $e^{-\frac{1}{2}xx'}$ 在过原点两平面间的积分,与这两平面的夹角成比例.立刻可有以下的推广,求过原点三平面间积分,我们可以猜测一定与“立体角”成比例,这是事实,切实些说.

$$\int_{\substack{-\infty \\ ux'>0 \\ vx'>0 \\ wx'>0}}^{\infty}\cdots\int e^{-\frac{1}{2}xx'}dx_1\cdots dx_n=\frac{1}{2}(2\pi)^{\frac{n}{2}-1}(\alpha+\beta+\gamma-\pi),$$

这儿 α, β, γ 是这三矢量的夹角,更一般些.

$$\int_{\substack{-\infty \\ ux'>0 \\ vx'>0 \\ wx'>0}}^{\infty}\cdots\int e^{-\frac{1}{2}xSx'}dx_1\cdots dx_n=\frac{1}{2}(2\pi)^{\frac{n}{2}-1}|S|^{-1/2}\left(\cos^{-1}\frac{-uS^{-1}v'}{\sqrt{uS^{-1}u'vS^{-1}v'}}\right.$$

$$\left.+\cos^{-1}\frac{-vS^{-1}w'}{\sqrt{vS^{-1}v'wS^{-1}w'}}+\cos^{-1}\frac{-wS^{-1}u'}{\sqrt{wS^{-1}w'uS^{-1}u'}}-\pi\right).$$

§ 4. 正态 Parent 分布

在研究正态 Parent 分布时用到以下的积分，对称方阵 $X=(x_{ij})$ 可以看成为 $\frac{1}{2}n(n+1)$ 维空间的一点. 所有的定正方阵 X 所成的区域以 $X>0$ 表之.

定理 1 假定 $k>n$, 对任一 $A>0$ 常有

$$\int_{X>0}\cdots\int |X|^{\frac{1}{2}(k-n-2)} e^{-\sigma(Ax)}\dot{X}$$
$$=\pi^{\frac{1}{4}n(n-1)}\Gamma\left(\frac{1}{2}(k-1)\right)\Gamma\left(\frac{1}{2}(k-2)\right)\cdots\Gamma\left(\frac{1}{2}(k-n)\right)|A|^{-\frac{1}{2}(k-1)},$$

此处 $\dot{X}=\prod_{1\leqslant i\leqslant j\leqslant n} dx_{ij}$。

证明　命 $A=TT'$，命 $T'XT=Y$，则由

$$|T|^{n+1}\dot{X}=\dot{Y}$$

(证明见 § 5) 可知

$$\int_{X>0}\cdots\int |X|^{\frac{1}{2}(k-n-2)} e^{-\sigma(AX)}\dot{X}=\int_{Y>0}\cdots\int |Y|^{\frac{1}{2}(k-n-2)}|T|^{-(k-n-2)}\times e^{-\sigma}(Y)|T|^{-(n+1)}\dot{Y}$$
$$=|A|^{-\frac{1}{2}(k-1)}\int_{Y>0}\cdots\int |Y|^{\frac{1}{2}(k-n-2)}e^{-\sigma(Y)}\dot{Y}.$$

即不妨假定 $A=I$ 时进行研究．即待证：命

$$\int_{X>0}\cdots\int |X|^{\frac{1}{2}(k-n-2)}e^{-\sigma(X)}\dot{X}=C_{n,k},$$

则

$$C_{n,k}=\pi^{\frac{1}{4}n(n-1)}\Gamma\left(\frac{1}{2}(k-1)\right)\Gamma\left(\frac{1}{2}(k-2)\right)\cdots\Gamma\left(\frac{1}{2}(k-n)\right).$$

换变数 $x_{ij}=y_{ij}\sqrt{x_{ii}x_{jj}}$，命

$$Y=\begin{pmatrix}1, & y_{12}, & \cdots, & y_{1n}\\ y_{21}, & 1, & \cdots, & y_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ y_{n1}, & y_{n2}, & \cdots, & 1\end{pmatrix},\quad y_{ij}=y_{ji},$$

则

$$X=DYD,\quad D=[\sqrt{x_{11}},\sqrt{x_{22}},\cdots,\sqrt{x_{nn}}].$$

由此得 $|X|=|Y|x_{11}\cdots x_{nn}$，而 Jacobian 等于

$$(x_{11}\cdots x_{nn})^{\frac{1}{2}(n-1)}.$$

因此

$$C_{k,n}=\int_0^\infty\cdots\int_0^\infty (x_{11}\cdots x_{nn})^{\frac{1}{2}(k-3)}e^{-\sum_{i=1}^n x_{ii}}dx_{11}\cdots dx_{nn}$$
$$\times\int_{Y>0}\cdots\int |Y|^{\frac{1}{2}(n-k-2)}dy_{12}\cdots dy_{n-1n}$$
$$=\Gamma\left(\frac{1}{2}(k-1)\right)^n J_n.$$

此处

$$J_n=\int_{Y>0}\cdots\int |Y|^{\frac{1}{2}(k-n-2)}dy_{12}\cdots dy_{n-1,n}.\quad(2\leqslant n<k).$$

用归纳法来算出此积分的数值．显然有

$$J_2=\int_{-1}^{1}(1-y^2)^{\frac{1}{2}(k-4)}dy=\sqrt{\pi}\frac{\Gamma\left(\frac{1}{2}(k-2)\right)}{\Gamma\left(\frac{1}{2}(k-1)\right)}.$$

用行列式的展开可知

$$\begin{vmatrix}1, & y_{12} & \cdots, & y_{1n+1}\\ \cdots & \cdots & \cdots & \cdots\\ y_{n+1,1}, & \cdots & \cdots, & 1\end{vmatrix}=|Y|-\sum_{ij=1}^{n}Y_{ij}y_{i,n+1}y_{j,n+1}.$$

此处 Y_{ij} 表示 y_{ij} 在 Y 中的余子式，因此

$$J_{n+1}=\int\cdots\int dy_{12}\cdots dy_{n-1,n}\int\left(|Y|-\sum_{i,j=1}^{n}Y_{ij}y_{in+1}y_{jn+1}\right)^{\frac{1}{2}(k-n-3)}dy_{1,n+1}\cdots dy_{n,n+1}.$$

此处内积分过所有的适合于

$$|Y|-\sum_{i,j=1}^{n}Y_{ij}y_{i,n+1}y_{j,n+1}>0$$

的 $y_{1,n+1},\cdots,y_{n,n+1}$．这个积分是可以用 Dirichlet 积分算出的，从而得出

$$J_{n+1}=J_n\frac{\Gamma\left(\frac{1}{2}(k-n-1)\right)}{\Gamma\left(\frac{1}{2}(k-1)\right)}\pi^{\frac{n}{2}}.$$

因此得出本定理．

由于定理 1 中表达出来的是 A 的整函数．把 a_{ij} 换为 $a_{ij}-i\varepsilon_{ij}t_{ij}$ 也对，此处 $\varepsilon_{ij}=1$ 若 $i=j$，而 $\varepsilon_{ij}=\frac{1}{2}$ 若 $i\neq j$．因而得出：

定理 2 命

$$f_k(X)=\begin{cases}C_{n,k}|A|^{\frac{1}{2}(k-1)}|X|^{\frac{1}{2}(k-n-2)}e^{-\sigma(AX)}, & X>0,\\ 0, & \text{不然}.\end{cases}$$

则 $f_k(X)$ 的 Fourier transform 等于

$$\varphi_n(T)=\int\cdots\int f_k(X)e^{i\sum_{i\leqslant j}x_{ij}t_{ij}}f_k(X)\cdot dx_{11}dx_{12}\cdots dx_{nn}=\left(\frac{|A|}{|A-iT|}\right)^{\frac{1}{2}(k-1)}.$$

这儿 $T=(\varepsilon_{ij}t_{ij})$.

§5. 矩阵变换的行列式

一个 $m\times n$ 矩阵可以看成为 mn 维空间的一点．两个 $m\times n$ 矩阵 X, Y 间的关系

$$X=PYQ,\ P=P^{(m)},\ Q=Q^{(n)} \tag{1}$$

可以看成为这个 mn 维空间的一个线性变换，现在算出这个线性变换的行列式．这个变换是由以下的两个变换而合成的

$$X=ZQ \tag{2}$$

$$Z=PY. \tag{3}$$

把 (2) 分行写出，得

$$(x_{11},\cdots,x_{1n})=(z_{11},\cdots,z_{1n})Q,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$(y_{m1},\cdots,x_{mn})=(z_{m1},\cdots,z_{mn})Q.$$

因此,线性变换 (2) 的行列式等于 $|Q|^m$. 同法,将 (3) 分列写出,得线性变换 (3) 的行列式等于 $|P|^n$. 因此(1)所定义的线性变换的行列式是

$$|P|^n|Q|^m. \tag{4}$$

特别当 $m=n$ 时,变换

$$X=PY\bar{P}'$$

的行列式等于 $|P\bar{P}'|^n=\text{abs}|P|^{2n}$.

再考虑 Hermitian 方阵,一个 Hernrtian 方阵

$$H=(h_{ij}),\quad h_{ii}=k_{ii},\quad h_{ij}=k_{ij}+ik_{ji},\quad i<j, \tag{5}$$

可以看成为 n^2 维空间的一点. 从 h_{ij} 变为 k_{ij} 的变形

$$h_{ii}=k_{ii},\quad \begin{matrix} h_{ij}=k_{ij}+ik_{ji}, \\ h_{ji}=k_{ij}-ik_{ji}, \end{matrix}\quad i<j,$$

的行列式等于

$$\begin{vmatrix} 1 & i \\ 1 & -i \end{vmatrix}^{\frac{1}{2}n(n-1)}=(-2i)^{\frac{1}{2}n(n-1)}.$$

因此由 Hermitian 关系

$$H_1=PH\bar{P}' \tag{6}$$

所定义出的线性变换的行列式等于

$$\text{abs}|P|^{2n}. \tag{7}$$

又 n 行列的对称方阵 S 成一 $\frac{1}{2}n(n+1)$ 维的空间. 而由

$$S=PTP' \tag{8}$$

所得的 S 与 T 的元素间的关系也成一线性变换,这变换的行列式等于

$$|P|^{n+1}. \tag{9}$$

这结果的证明简述如次,如果 P 是平延,不难证明所求的行列式的值等于 1. 如果

$$P=[\lambda,1,\cdots,1],$$

则由

$$s_{11}=\lambda^2 t_{11},\ s_{12}=\lambda t_{12},\cdots,s_{1n}=\lambda t_{1n},$$

(其它不变)可得所求的行列式之值等于 λ^{n+1}. 因而得出本结论.

斜对称方阵 K 成一 $\frac{1}{2}n(n-1)$ 维空间. 由

$$K=PQP' \tag{10}$$

所得的线性变换的行列式等于

$$|P|^{n-1}. \tag{11}$$

读者自证之,(读者考虑由 $X=PYP'$ 也推出

$$X+X'=P(Y+Y')P',\ X-X'=P(Y-Y')P',$$

因而看出 (9) 乘 (11) 等于$|P|^{2n}$的道理).

附记 1) 把(1)中的X的元素排成为

$$x_{11}, \cdots, x_{1m}, x_{21}, \cdots, x_{2m}, \cdots, x_{m1}, \cdots, x_{mn},$$

同样排好 Y 的元素,所得出的方阵用

$$P' \cdot \times Q$$

来表之,而 $R \cdot \times Q$ 定义为 mn 行列的方阵

$$\begin{pmatrix} r_{11}Q, & \cdots, & r_{1m}Q \\ \cdots & \cdots & \cdots \\ r_{m1}Q, & \cdots, & r_{mm}Q \end{pmatrix}$$

称为R与Q的直乘积. 公式 (4) 也可以写成为 $|P' \cdot \times Q| = |P|^n |Q|^m$.

2) 把 (8) 中 S 的元素写成为

$$S_{11}, S_{12}, \cdots, S_{1n}, S_{22}, S_{23}, \cdots, S_{2n}, \cdots, S_{n-1n}, S_{nn}$$

并相仿地排好T的元素. 这样 (8) 所定义的方阵用$P^{[2]}$表之. (9) 说明了

$$|P^{[2]}| = |P|^{n+1}.$$

同样由斜对称定义所得的方阵用$P^{(2)}$表之,(11) 说明了

$$|P^{(2)}| = |P|^{n-1}.$$

3) $P' \cdot \times P, P^{[2]}, P^{(2)}$ 有次之性质

$$P^{[2]}Q^{(2)} = (PQ)^{[2]},\ P^{(2)}Q^{(2)} = (PQ)^{(2)}$$

及

$$(P' \cdot \times P)(Q' \cdot \times Q) = (PQ)' \cdot \times PQ.$$

这些性质是群表示论中最原始的例子.

§ 6. 酉群上的积分元素

把 n 行列的复元素方阵 Z 看成为 $2n^2$ 维实空间的一点,并且有 Euclid 度量

$$\sum_{i=1}^{n} \sum_{j=1}^{n} (dx_{ij}^2 + dy_{ij}^2),\quad z_{ij} = x_{ij} + iy_{ij}. \tag{1}$$

以上的 Euclid 度量也可以写成为

$$\sigma(dZ\,\overline{dZ'}) \tag{2}$$

的形式,这儿$\sigma(A)$表示方阵A之迹.

所有的酉方阵成为这空间的一个流形,现在先证明它的维数是 n^2,作 Cayley 变换

$$U = (I + iH)^{-1}(I - iH), \tag{3}$$

可以解得

$$iH = (I - U)(I + U)^{-1}. \tag{4}$$

酉方阵所适合的条件一变而为

$$I = U\overline{U}' = (I + iH)^{-1}(I - iH)(I + i\overline{H}')(I - i\overline{H}')^{-1}$$

即

$$(I - iH)(I + i\overline{H}') = (I + iH)(I - i\overline{H}'),$$

即

$$H=\overline{H}'. \tag{5}$$

由此可知，一个使$|I+U|\neq 0$的酉方阵对应于一个 Hermite 方阵H. H中有n^2个自变数，因此酉方阵所成的流形是n^2维的，但须注意，适合$|I+U|=0$的酉方阵成一较低维的流形.

H可以看成酉方阵所成流形的参变数. 反过来，酉群流形也可以看为“Hermitian”方阵所成的空间的扩张空间，即加上一无穷远点的流形所成的空间.

酉群流形上的度量当然是

$$\sigma(dU\overline{dU'}).$$

微分

$$U\overline{U}'=I$$

得

$$dU\overline{U}'+Ud\overline{U}'=0.$$

命 $\delta U=U^{-1}dU$，则 $\delta U=-\overline{\delta U'}$. 因而

$$\sigma(dUd\overline{U}')=-\sigma(\delta U\overline{\delta U'}). \tag{6}$$

另一方面由 Cayley 表达式

$$\begin{aligned}dU&=-(I+iH)^{-1}idH-i(I+iH)^{-1}dH(I+iH)^{-1}(I-iH)\\&=-i(I+iH)^{-1}dH[I+(I+iH)^{-1}(I-iH)]\\&=-2i(I+iH)^{-1}dH(I+iH)^{-1}.\end{aligned} \tag{7}$$

由此得出

$$\begin{aligned}\sigma(dUd\overline{U}')&=4\sigma((I+iH)^{-1}dH(I+iH)^{-1}\times(I-iH)^{-1}dH(I-iH)^{-1})\\&=4\sigma(dH(I+H^2)^{-1}dH(I+H^2)^{-1}).\end{aligned} \tag{8}$$

命 $(I+H^2)^{-1}=P\overline{P}'$ 及 $X=\overline{P}'dHP$，则得

$$\sigma(dUd\overline{U}')=4\sigma(X^2)=4\sum_{i,j=1}^{n}|x_{ij}|^2.$$

由于 $\bar{x}_{ij}=x_{ji}$，命 $x_{ij}=y_{ij}+iz_{ij}$，$|x_{ij}|^2+|x_{ji}|^2=2(y_{ij}^2+z_{ij}^2)$，因此二次型

$$4\sum_{i,j=1}^{n}|x_{ij}|^2$$

的行列式等于

$$4^{n^2}\cdot 2^{n(n-1)}=2^{n(3n-1)}.$$

由于 $X=\overline{P}'dHP$ 及上节(6)，(7)可知由X变为dH的线性变换的行列式等于

$$|P\overline{P}'|^n=|I+H^2|^{-n}.$$

因此，(8)式的行列式等于

$$2^{n(3n-1)}\cdot|I+H^2|^{-2n}.$$

因此酉群上的体积元素等于

$$\dot{U}=2^{n^2}\cdot 2^{\frac{1}{2}n(n-1)}|I+H^2|^{-n}\prod_{j=1}^{n}dh_j\prod_{j<k}dh'_{jk}dh''_{jk}.$$

这儿 $H=(h_{jk})$，$h_{jj}=h_j$，$h_{jk}=h'_{jk}+ih''_{jk}$，整个酉群的积分范围是

$$-\infty<h_j<\infty,\quad -\infty<h'_{jk},\ h''_{jk}<\infty.$$

定理 1 酉群的体积等于

$$\vartheta_n = \frac{(2\pi)^{\frac{1}{2}n(n+1)}}{1!2!\cdots(n-1)!}.$$

由以上的推导可知

$$\vartheta_n = \int\cdots\int_U \dot{U} = 2^{n^2}\cdot 2^{\frac{1}{2}n(n-1)}\int_{-\infty}^{\infty}\cdots\int |I + H^2|^{-n}\prod_{j=1}^{n} dh_j \prod_{j<k} dh'_{jk}dh''_{jk}$$

这个积分的算出见下节。

§7. （续）

在算出上节所要求的积分之前，先算出以下的定积分。

定理 1 命 $a > 0$，$b^2 - ac < 0$，$\alpha > \frac{1}{2}$，则

$$\int_{-\infty}^{\infty}\frac{dx}{(ax^2+2bx+c)^\alpha} = a^{\alpha-1}(ac-b^2)^{\frac{1}{2}-\alpha}\sqrt{\pi}\frac{\Gamma\left(\alpha-\frac{1}{2}\right)}{\Gamma(v)}.$$

证明　命

$$y = \frac{a}{\sqrt{ac-b^2}}\left(x+\frac{b}{a}\right),$$

则

$$dx = \frac{\sqrt{ac-b^2}}{a}dy$$

及

$$ax^2+2bx+c = \frac{ac-b^2}{a}(y^2+1).$$

于是

$$\int_{-\infty}^{\infty}\frac{dx}{(ax^2+2bx+c)^\alpha} = a^{\alpha-1}(ac-b^2)^{\frac{1}{2}-\alpha}\int_{-\infty}^{\infty}\frac{dy}{(1+y^2)^\alpha}$$

$$= a^{\alpha-1}(ac-b^2)^{\frac{1}{2}-\alpha}\sqrt{\pi}\frac{\Gamma\left(\alpha-\frac{1}{2}\right)}{\Gamma(\alpha)}.$$

定理 2 设 $\alpha > n - \frac{1}{2}$，H 表 n 行列的 Hermite 方阵，则

$$H_n(\alpha) = \int\cdots\int_H \frac{\dot{H}}{(\det(I+H^2))^\alpha}$$

$$= 2^{\frac{n(n-1)}{2}}\pi^{\frac{1}{2}n^2}\prod_{i=0}^{n-1}\frac{\Gamma\left(\alpha-j-\frac{1}{2}\right)}{\Gamma(\alpha-j)}\prod_{k=0}^{n-2}\frac{\Gamma(2\alpha-n-k)}{\Gamma(2\alpha-2k-1)},$$

此处 $H = (h_{jk})$，$h_{jj} = h_j$，$h_{jk} = h'_{jk} + ih''_{jk}(j<k)$，

$$\dot{H} = 2^{\frac{n(n-1)}{2}}\prod_{i=1}^{n} dh_i \prod_{j<k} dh'_{jk}dh''_{jk}.$$

证明　命

$$H=\begin{pmatrix}H_1 & \bar{v}'\\ v & h\end{pmatrix},\quad (h=h_n)$$

其中 H_1 为一 $n-1$ 行列的 Hermite 方阵，v 为一 $n-1$ 维矢量，h 为一实数，则

$$I+H^2=\begin{pmatrix}I+H_1^2+\bar{v}'v & H_1\bar{v}'+\bar{v}'h\\ vH_1+hv & 1+h^2+v\bar{v}'\end{pmatrix}.$$

利用等式

$$\begin{pmatrix}I & 0\\ -pA^{-1} & 1\end{pmatrix}\begin{pmatrix}A & \bar{p}'\\ p & l\end{pmatrix}\overline{\begin{pmatrix}I & 0\\ -pA^{-1} & 1\end{pmatrix}}'=\begin{pmatrix}A & 0\\ 0 & l-pA^{-1}\bar{p}'\end{pmatrix},\quad (A=\bar{A}')$$

可得

$$\det(I+H^2)=(ah^2+2bh+c)\det(I+H_1^2+\bar{v}'v),$$

此处

$$a=1-v(I+H_1^2+\bar{v}'v)^{-1}\bar{v}',$$
$$2b=-vH_1(I+H_1^2+\bar{v}'v)^{-1}\bar{v}'-v(I+H_1^2+\bar{v}'v)^{-1}H_1\bar{v}',$$
$$c=1+v\bar{v}'-vH_1(I+H_1^2+\bar{v}'v)^{-1}H_1\bar{v}'.$$

因 H_1 为 Hermite 方阵，故存在一酉方阵 U 使

$$H_1=U[\lambda_1,\lambda_2,\cdots,\lambda_n]\bar{U}'.$$

命

$$T\doteq U[\sqrt{1+\lambda_1^2},\sqrt{1+\lambda_2^2},\cdots,\sqrt{1+\lambda_n^2}]\bar{U}',$$

则

$$T=\bar{T}',\quad TH_1=H_1T,\quad I+H_1^2=T^2.$$

再作变换

$$v=uT,$$

则

$$\dot{v}=|\det T|^2\dot{u}=\det(I+H_1^2)\dot{u},$$
$$I+H_1^2+\bar{v}'v=T(I+\bar{u}'u)T.$$

又因为

$$(I+\bar{u}'u)^{-1}\bar{u}'=\frac{\bar{u}'}{1+u\bar{u}'},\quad w(I+\bar{u}'u)^{-1}\bar{w}'=w\bar{w}'-\frac{|w\bar{u}'|^2}{1+u\bar{u}'},$$

此处 w 为一 n 维矢量，故可知

$$a=1-u(I+\bar{u}'u)^{-1}\bar{u}'=\frac{1}{1+u\bar{u}'}\quad(>0),$$
$$b=-uH_1(I+\bar{u}'u)^{-1}\bar{u}'=-\frac{uH_1\bar{u}'}{1+u\bar{u}'},$$
$$c=1+uT^2\bar{u}'-uH_1(I+\bar{u}'u)^{-1}H_1\bar{u}'=1+u\bar{u}'+\frac{|uH_1\bar{u}'|^2}{1+u\bar{u}'}.$$

由于 $uH_1\bar{u}'$ 是实数，故得

$$ac-b^2=1.$$

由定理 1 可得

$$
\begin{aligned}
H_n(\alpha) &= \int_H\cdots\int \frac{\dot{H}}{(\det(I+H^2))^{\alpha}} \\
&= 2^{n-1}\int_{u,H_1}\cdots\int (\det(I+H_1^2))^{1-\alpha}(1+u\bar{u}')^{-\alpha}\dot{u}\dot{H}_1\int_{-\infty}^{\infty}(ah^2+2bh+c)^{-\alpha}dh \\
&= 2^{n-1}\frac{\pi^{\frac{1}{2}}\Gamma\left(\alpha-\frac{1}{2}\right)}{\Gamma(\alpha)}\int_u\cdots\int(1+u\bar{u}')^{1-2\alpha}\dot{u}\int_{H_1}\cdots\int(\det(I+H_1^2))^{1-\alpha}\dot{H}_1 \\
&= 2^{n-1}\pi^{n-\frac{1}{2}}\frac{\Gamma\left(\alpha-\frac{1}{2}\right)\Gamma(2\alpha-n)}{\Gamma(\alpha)\Gamma(2\alpha-1)}H_{n-1}(\alpha-1).
\end{aligned}
$$

继续应用此式，并直接算出

$$
H_1(\alpha-n+1)=\int_{-\infty}^{\infty}\frac{dx}{(1+x^2)^{\alpha-n+1}}=\frac{\pi^{\frac{1}{2}}\Gamma\left(\alpha-n+\frac{1}{2}\right)}{\Gamma(\alpha-n+1)},\ \left(\alpha>n-\frac{1}{2}\right)
$$

即得定理.

附记 1) 上节所定义的酉积分元素 $\dot{U}$ 有它的不变性，即如果

$$
U_1=VUW,\quad V,W\text{是酉方阵},
$$

则

$$
\sigma(dU_1\overline{dU_1'})=\sigma(dU\overline{dU'}).
$$

因而 $\dot{U}_1=\dot{U}$.

这是不变积分的一个例子.

§8. 实正交方阵的体积元素

我们现在叙述一下实正交群的积分元素及其体积.

现在先研究 n 行列的实正交群 O_n，即适合以下关系的 n 行列的实方阵 T:

$$
T'T=I. \tag{1}
$$

显然有 $\det T=\pm1$. 行列式为 $+1$ 的正交方阵所成的群用 O_n^+ 表示. 对应于一个 T 可以做一个方阵

$$
K=(I-T)(I+T)^{-1}, \tag{2}
$$

但 $\det(I+T)=0$ 的情形必须除外. 现在对 $\det(I+T)=0$ 的情形我们不能说"一般说来不成立". 因为任一行列式为 -1 的 T，一定使 $\det(I+T)=0$ (此点可由 $\det(I+T)=\det(TT'+T)=\det T\det(T'+I)=-\det(I+T)$ 得知)，所以现在限定 T 属于 O_n^+. 由 (1) 立得

$$
K=-K'; \tag{3}
$$

解 (2) 立得

$$
T=(I-K)(I+K)^{-1}; \tag{4}
$$

由于 $\det(I-K)=\det(I-K')=\det(I+K)$，也可知 $\det T=+1$.

微分 (4) 可得

$$dT = -2(I+K)^{-1}dK(I+K)^{-1},$$

因此

$$\sigma(dTdT') = -4\sigma(dK(I-K^2)^{-1}dK(I-K^2)^{-1}). \tag{5}$$

把 dK 写成 (dk_{ij})，其中 $dk_{ij} = -dk_{ji}$；把 $(I-K^2)^{-1}$ 写成 (u_{st})，其中 $u_{st} = u_{ts}$，则 (5) 变成为

$$8\sum_{i<j}\sum_{s<t}(u_{js}u_{it} - u_{is}u_{jt})dk_{ij}dk_{st},$$

故得出积分元素

$$\dot{T} = 2^{\frac{1}{2}n(n-1)}\det(I-K^2)^{-\frac{1}{2}(n-1)}\dot{K}, \tag{6}$$

此处 $\dot{K} = 2^{\frac{n(n-1)}{4}}\prod_{i<j}dk_{ij}$.

§9. 实正交群的总体积

定理 1 命 $n \geqslant 2$. 若 $\alpha > \frac{1}{4}(2n-3)$，则

$$J_n(\alpha) = \int\cdots\int_K \frac{\dot{K}}{(\det(I+KK'))^{\alpha}}$$

$$= 2^{\frac{1}{4}n(n-1)}\pi^{\frac{1}{4}n(n-1)}\prod_{\nu=2}^{n}\frac{\Gamma\left(2\alpha - n + \frac{1}{2}(\nu+1)\right)}{\Gamma(2\alpha - n + \nu)}.$$

此处K过所有的 n 行列的实斜对称方阵，$K = (k_{ij})$，$\dot{K} = 2^{\frac{n(n-1)}{4}}\prod_{i<j}dk_{ij}$.

证明　把K写成

$$K = \begin{pmatrix} K_1 & -t' \\ t & 0 \end{pmatrix},$$

此处 K_1 为一 $n-1$ 行列的实斜对称方阵，t 为一 $n-1$ 维矢量，于是

$$I + KK' = \begin{pmatrix} I + K_1K_1' + t't & K_1t' \\ tK_1' & 1 + tt' \end{pmatrix}.$$

不难证明

$$\det(I+KK') = (1 + tt' - tK_1'(I + K_1K_1' + t't)^{-1}K_1t')\det(I + K_1K_1' + t't).$$

有一正交方阵Γ使

$$K_1 = \Gamma\left(\begin{pmatrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{pmatrix} \dotplus \begin{pmatrix} 0 & \lambda_2 \\ -\lambda_2 & 0 \end{pmatrix} \dotplus \cdots\right)\Gamma',$$

括弧中最后一项或为 $\begin{pmatrix} 0 & \lambda_{\frac{n}{2}} \\ -\lambda_{\frac{n}{2}} & 0 \end{pmatrix}$ 或为 $\begin{pmatrix} 0 & \lambda_{\frac{n-1}{2}} \\ -\lambda_{\frac{n-1}{2}} & 0 \end{pmatrix} \dotplus 0$，各视 n 为偶或奇而定.

命

$$T = \Gamma[\sqrt{1+\lambda_1^2}, \sqrt{1+\lambda_1^2}, \sqrt{1+\lambda_2^2}, \sqrt{1+\lambda_2^2}, \cdots]\Gamma',$$

则

$$T = T',\ K_1T = TK_1,\ I + K_1K_1' = T^2.$$

又命 $t=wT$，则

$$\dot{t}=(\det T)\dot{w}=(\det(I+K_1K_1'))^{\frac{1}{2}}\dot{w},$$
$$I+K_1K_1'+t't=T^2+Tw'wT=T(I+w'w)T.$$

所以

$$\begin{aligned}1+tt'-tK_1'(I+K_1K_1'+t't)^{-1}K_1t'\\=1+wT^2w'-wTK_1'T^{-1}(I+w'w)^{-1}T^{-1}K_1Tw'\\=1+wT^2w'-wK_1'(I+w'w)^{-1}K_1w'\\=1+ww'-\frac{(wK_1w')^2}{1+ww'}=1+ww'.\end{aligned}$$

于是

$$\begin{aligned}J_n(\alpha)&=\int\cdots\int_K\frac{\dot{K}}{(\det(I+KK'))^{\alpha}}\\&=2^{\frac{n-1}{2}}\int\cdots\int_{K_1}\frac{\dot{K}_1}{\det(I+K_1K_1')^{\alpha-\frac{1}{2}}}\int\cdots\int_w(1+ww')^{-2\alpha}\dot{w}\\&=2^{\frac{n-1}{2}}\pi^{\frac{1}{2}(n-1)}\frac{\Gamma\left(2\alpha-\frac{1}{2}(n-1)\right)}{\Gamma(2\alpha)}J_{n-1}\left(\alpha-\frac{1}{2}\right).\end{aligned}$$

连续应用此式并最后算出

$$J_2\left(\alpha-\frac{n-2}{2}\right)=\sqrt{2}\int_{-\infty}^{\infty}\frac{dt}{(1+t^2)^{2\alpha-n+2}}=\sqrt{2}\pi^{\frac{1}{2}}\frac{\Gamma\left(2\alpha-n+\frac{3}{2}\right)}{\Gamma(2\alpha-n+2)},$$
$$\left(\alpha>\frac{1}{4}(2n-3)\right)$$

即得定理．所以有

定理 2 正交群 O_n^+ 的总体积是

$$\begin{aligned}\int\dot{T}&=2^{\frac{1}{2}n(n-1)}\int\cdots\int_K\det(I-K^2)^{-\frac{1}{2}(n-1)}\dot{K}\\&=2^{\frac{3}{4}n(n-1)}\pi^{\frac{1}{4}n(n-1)}\prod_{\nu=2}^{n}\frac{\Gamma\left(\frac{1}{2}(\nu-1)\right)}{\Gamma(\nu-1)}.\end{aligned}$$

习题 1 试算出辛酉群的总体积．

习题 2 试算出对称酉方阵所成的流形的总体积．

习题 3 试算出斜对称酉方阵所成的流形的总体积．

第九章　非负方阵

§1. 非负方阵的相似性

定义 1　如果一个方阵 Q，其中每行每列只有一个正元素，其他的元素都是 0，则称 Q 为广义换位方阵.

例如；对角线方阵 $\Lambda=[\lambda_1,\cdots\lambda_n]$ $(\lambda_i>0)$ 即为广义换位方阵，又如换位方阵 P，它的每行中与每列中只有一个元素为 1，其他元素都是 0，也是广义换位方阵，$P\Lambda$ 也是广义换位方阵. 不难看出，此外没有其他广义换位方阵了.

定义 2　命 A，B 是两个非负方阵，如果有一个广义换位方阵 Q 使

$$QAQ^{-1}=B,$$

则 A 与 B 称为相似，用符号 $A\sim B$ 表它.

显然有以下的性质.

1) $A\sim A$.

2) 如果 $A\sim B$，则 $B\sim A$.

3) 如果 $A\sim B$，$B\sim C$，则 $A\sim C$.

引进符号 $A\geqslant 0$，表示矩阵 A 的元素都是非负的而 $A>0$ 表示 A 的元素都是正的，同样定义矢量 $x\geqslant 0$ 及 $x>0$. 进一步引伸，$A\geqslant B$ 表示 $A-B\geqslant 0$ 等等. 如此则显然有

4) $A\geqslant B$，$B\geqslant C$ 则 $A\geqslant C$.

5) $A\geqslant 0$，$B\geqslant C$ 则 $AB\geqslant AC$ 及 $BA\geqslant CA$.

6) 如果 $A\sim B$，$A\geqslant 0$ (或>0)，则 $B\geqslant 0$ (或>0).

这一性质的证明依赖于广义换位方阵和其逆方阵都是非负的. 反之，有

定理 1　如果一个可逆非负方阵的逆方阵也是非负的，那末它一定是广义换位方阵.

证明　命

$$B=(b_{ij}),\quad b_{ij}\geqslant 0$$

是

$$A=(a_{ij}),\quad a_{ij}\geqslant 0$$

的逆方阵，则当 $i\neq k$ 时，

$$\sum_{j=1}^{n}a_{ij}b_{jk}=0.$$

由此得出，当 $i\neq k$ 时，对任一 j 常有

$$a_{ij}b_{jk}=0.$$

如果 A 的第 i 行中有两个元素 $a_{ij_1}\neq 0$，$a_{ij_2}\neq 0$，则对所有的 $k(\neq i)$ 常有

$$b_{j_1k}=b_{j_2k}=0,$$

也就是 B 的第 j_1，j_2 行中除去 b_{j_1i}，b_{j_2i} 二元素外都等于 0，这样的 B 是奇异的，与假定相

矛盾，因此A中每一行只有一个非零元素，同样每一列中亦只有一个非零元素，定理证完.

习题 1　对任何非负方阵A，

$$TAT^{-1}$$

仍是非负方阵，则T一定是广义换位方阵.

定义 3　如果

$$A \sim \begin{pmatrix} A_1^{(h)} & A_2 \\ 0 & A_3^{(n-h)} \end{pmatrix},$$

则A称为可分拆的，不然称为不可分拆的.

§2. 标　准　型

我们现在讨论不可分拆的非负方阵.

定义 1　一个不可分拆的非负方阵A如果适合于

$$A = (a_{ij}), \quad \sum_{j=1}^{n} a_{ij} = q,$$

则称为标准型，而q称为高标.

基本定理　任何一个不可拆的非负方阵一定相似于一个标准型方阵. 如果不计其重排所得的情况，标准型是唯一的.

在证明这定理之前，先介绍一个方法，这个方法是证明的源泉，但也是进行计算的好方法.

假定

$$\begin{cases} a_{11} + \cdots + a_{1n} = q_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1} + \cdots + a_{nn} = q_n. \end{cases} \tag{1}$$

及 $q_1 \geqslant q_2 \geqslant \cdots \geqslant q_n$，而且 $q_1 > q_n$.

取λ使

$$a_{11} + \cdots + a_{1n-1} + \frac{a_{1n}}{\lambda} = \lambda(a_{n1} + \cdots + a_{nn-1}) + a_{nn}. \tag{2}$$

$$q_1 - a_{1n} + \frac{a_{1n}}{\lambda} = \lambda(q_n - a_{nn}) + a_{nn}.$$

即

$$f(\lambda) = \lambda^2(q_n - a_{nn}) + (a_{nn} + a_{1n} - q_1)\lambda - a_{1n} = 0.$$

如果 $q_n = a_{nn}$，则 $a_{n1} = \cdots = a_{nn-1} = 0$，$A$是可分拆的方阵了，所以 $q_n > a_{nn}$. 由于 $f(\pm\infty) > 0$，$f(1) = q_n - q_1 < 0$. 所以有 $\lambda > 1$ 适合于(2). 由于 $\lambda > 1$，所以方阵

$$[1,\cdots,1,\lambda]A[1,\cdots,1,\lambda]^{-1} = \begin{pmatrix} a_{11}, & a_{12}, & \cdots & a_{1n-1}, & \dfrac{a_{1n}}{\lambda} \\ a_{21}, & a_{22}, & \cdots & a_{2n-1}, & \dfrac{a_{2n}}{\lambda} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda a_{n1}, & \lambda a_{n2}, & \cdots & \lambda a_{nn-1}, & a_{nn} \end{pmatrix}$$

的行和都$<q_1$(有例外，例如 $a_{2n}=0$，$q_2=q_1$)。

所以一般讲来，可以用这个办法来逐步减少 $\max(q_1,\cdots,q_n)$，一直到所有的行和都相等为止。

§3. 基本定理的证明

上节中所述及的计算方法，便是基本定理证明的根源，引进

$$Q(A)=\max(q_1,\cdots,q_n). \tag{1}$$

先证明如果 $q_1,\cdots,q_n$ 不全相等，我们可以取 Λ 使

$$Q(\Lambda A\Lambda^{-1})<Q(A). \tag{2}$$

假定经过重新排列，方阵 A 的行和 $q_1,\cdots,q_n$ 可以排成为

$$q_1=\cdots=q_s>q_{s+1}\geqslant q_{s+2}\geqslant\cdots\geqslant q_n.$$

我们证明可以取得 Λ 使 $\Lambda A\Lambda^{-1}$ 中和等于 q_1 的行数$<s$。其他的行的和都小于 q_1 把方阵 A 拆成为

$$A=\begin{pmatrix}A_1 & A_2\\ A_3 & A_4\end{pmatrix},\quad A_1=A_1^{(s)} \text{ 等等.}$$

取

$$\Lambda=\begin{pmatrix}I^{(s)} & 0\\ 0 & \Lambda_1\end{pmatrix},\quad \Lambda_1=[\lambda_{s+1},\cdots,\lambda_n],$$

如此得

$$\Lambda A\Lambda^{-1}=\begin{pmatrix}A_1 & A_2\Lambda_1^{-1}\\ \Lambda_1A_3 & \Lambda_1A_4\Lambda_1^{-1}\end{pmatrix}.$$

由于 $A_2\neq 0$，所以有 $\lambda_{s+1},\cdots,\lambda_n$ 使

$$A_2\Lambda_1^{-1}\leqslant A_2,\text{ 但 } A_2\Lambda_1^{-1}\neq A_2.$$

同时使 $\Lambda A\Lambda^{-1}$ 的后 $n-s$ 行的和仍然小于 q_1。由 $A_2\Lambda_1^{-1}\leqslant A_2$ 及 $A_2\Lambda_1^{-1}\neq A_2$ 可知 $\Lambda A\Lambda^{-1}$ 的前 s 行中至少有一行之和 $<q_1$，因而 $\Lambda A\Lambda^{-1}$ 中和等于 q_1 的行数$<s$。一直下去，一直到没有一行的和 $\geqslant q_1$。这证明了(2)式。

其次我们考虑集合 S：

$$\Lambda=[\lambda_1,\cdots,\lambda_n],\ \lambda_1+\cdots+\lambda_n=1,\ \lambda_\nu\geqslant 0. \tag{3}$$

$\left(\text{由于 } \Lambda A\Lambda^{-1}=\left(\frac{1}{\lambda}\Lambda\right)A\left(\frac{1}{\lambda}\Lambda\right)^{-1} \text{ 我们实质上已经讨论了所有的 } \Lambda \text{ 了}\right)$，对应于每一 $\Lambda(\in S)$，有一个数值

$$Q(\Lambda A\Lambda^{-1}).$$

命 q 代表这集合的确下界，如果有 Λ 使

$$Q(\Lambda A\Lambda^{-1})=q.$$

则由(2)可知，$\Lambda A\Lambda^{-1}$ 的各行之和都等于 q，即 A 相似于一个标准型。

由 q 的定义，在 S 中有一组方阵贯

$$\Lambda_1,\Lambda_2,\cdots,\Lambda_i,\cdots \tag{4}$$

使

$$Q(\Lambda_iA\Lambda_i^{-1})<q+\frac{1}{i}. \tag{5}$$

由于 S 是闭集，所以 $\{\Lambda_i\}$ 有一极限点 Λ_0，在(4)中可以选一子贯趋于 Λ_0。不妨假定

$$\lim_{i\to\infty}\Lambda_i=\Lambda_0.$$

先证明 $\Lambda_0=[\lambda_1^{(0)},\cdots,\lambda_n^{(0)}]>0$．如果有些 $\lambda_j^{(0)}=0$．但由 $\lambda_1^{(0)}+\cdots+\lambda_n^{(0)}=1$ 可知不能全部为 0．假定其中 $\lambda_1^{(0)}=\cdots=\lambda_s^{(0)}=0$，而 $\lambda_{s+1}^{(0)},\cdots,\lambda_n^{(0)}$ 都非零，由

$$\lambda_j^{(i)}a_{j1}\lambda_1^{(i)^{-1}}+\cdots+\lambda_j^{(i)}a_{js}\lambda_s^{(i)^{-1}}+\cdots+\lambda_j^{(i)}a_{jn}\lambda_n^{(i)^{-1}}<q+\frac{1}{i}$$

$$(s+1\leqslant j\leqslant n,\ i=1,2,\cdots)$$

可知 $a_{j1}=\cdots=a_{js}=0\,(s+1\leqslant j\leqslant n)$．即 A 是可分拆的了，此为矛盾，所以 $\Lambda_0>0$．

命 $\Lambda_i=[\lambda_1^{(i)},\cdots,\lambda_n^{(i)}]$．由 (5) 可知

$$\max[\lambda_1^{(i)}(a_{11}\lambda_1^{(i)^{-1}}+\cdots+a_{1n}\lambda_n^{(i)^{-1}}),\cdots,\lambda_n^{(i)}(a_{n1}\lambda_1^{(i)^{-1}}+\cdots+a_{nn}\lambda_n^{(i)^{-1}})]<q+\frac{1}{i}.$$

即对任一 j，有

$$\lambda_j^{(i)}(a_{j1}\lambda_1^{(i)^{-1}}+\cdots+a_{jn}\lambda_n^{(i)^{-1}})<q+\frac{1}{i},$$

命 $i\to\infty$．则得

$$\lambda_j^{(0)}(a_{j1}\lambda_1^{(0)^{-1}}+\cdots+a_{jn}\lambda_n^{(0)^{-1}})\leqslant q.$$

即

$$Q(\Lambda_0A\Lambda_0^{-1})=q.$$

即得定理的前半部.

最后证明唯一性，假定 A 是标准型，其行和是 q，而 $\Lambda A\Lambda^{-1}$ 也是标准型，行和是 q_1，则

$$a_{i1}\lambda_1^{-1}+\cdots+a_{in}\lambda_n^{-1}=q_1\lambda_i^{-1}.\qquad(6)$$

即对所有的 i 常有

$$q\min(\lambda_1^{-1},\cdots,\lambda_n^{-1})\leqslant q_1\lambda_i^{-1}\leqslant q\max(\lambda_1^{-1},\cdots,\lambda_n^{-1}),\qquad(7)$$

取 $\lambda_i^{-1}=\max(\lambda_1^{-1},\cdots,\lambda_n^{-1})$ 及 $\min(\lambda_1^{-1},\cdots,\lambda_n^{-1})$ 立得 $q=q_1$，代入(6)式得

$$a_{i1}\lambda_1^{-1}+\cdots+a_{in}\lambda_n^{-1}=(a_{i1}+\cdots+a_{in})\lambda_i^{-1}.\qquad(8)$$

如果重新排列成为

$$\lambda_1^{-1}\leqslant\cdots\leqslant\lambda_s^{-1}<\lambda_{s+1}^{-1}=\cdots=\lambda_n^{-1},$$

则由(8)可知当 $i=s+1,\cdots,n$ 时，

$$a_{i1}=\cdots=a_{is}=0.$$

这说明 A 是可分拆的，因此得出唯一性.

§4. 基本定理的另一形式

由

$$\Lambda A\Lambda^{-1}=(b_{ij}),\quad \sum_{j=1}^{n}b_{ij}=q$$

可知

$$\sum_{j=1}^{n}a_{ij}\frac{\lambda_i}{\lambda_j}=\sum_{j=1}^{n}b_{ij}=q,$$

即

$$\sum_{j=1}^{n}\frac{a_{ij}}{\lambda_j}=\frac{q}{\lambda_i}.$$

即得

$$A\begin{pmatrix}\lambda_1^{-1}\\ \vdots\\ \lambda_n^{-1}\end{pmatrix}=q\begin{pmatrix}\lambda_1^{-1}\\ \vdots\\ \lambda_n^{-1}\end{pmatrix}.$$

由此立刻推得:

定理 1 不可分拆的非负方阵，有一个正特征根 q，并且对应于它有一个正特征矢量(列).

关于正特征根 q 与正特征矢量我们还有以下一些性质.

定理 2 不可分拆的非负方阵只有一个非负特征(列)矢量(如果不计其常数因子的话)，它是正矢量而且对应的特征根是高标，其他的特征根的绝对值都不超过高标.

证明 不妨假定原来的方阵就是标准型，即

$$A=(a_{ij}),\quad \sum_{j=1}^{n}a_{ij}=q.$$

显然有$(1,1,\cdots,1)'$为其特征列矢量，它是正的. 若有另一非负特征列矢量$(x_1,\cdots,x_n)'$ $(\neq 0')$，则

$$a_{i1}x_1+\cdots+a_{in}x_n=q_1x_i,\ i=1,2,\cdots,n. \tag{1}$$

由此推出 $q_1>0$. 因而得

$$q\min(x_1,\cdots,x_n)\leqslant q_1x_i\leqslant q\max(x_1,\cdots,x_n). \tag{2}$$

如果 x_i 中有一些为 0，例如 $x_1=\cdots=x_s=0$，$x_{s+1}>0,\cdots,x_n>0$，则由

$$a_{is+1}x_{s+1}+\cdots+a_{in}x_n=q_1x_i=0,\ i=1,2,\cdots,s$$

可知 $a_{is+1}=\cdots=a_{in}=0$，$1\leqslant i\leqslant s$，即 A 是可分拆的，因此 x_i 都是正的.

由(2)推出 $q_1=q$，再由(1)得

$$a_{i1}x_1+\cdots+a_{in}x_n=(a_{i1}+\cdots+a_{in})x_i$$

可推得 $x_1=\cdots=x_n$. 因而得知非负特征矢量的唯一性，即只有一个非负特征矢量它对应于特征值为高标 q.

命 q_1 是 A 的任一特征根，它对应的列矢量是 $x'=(x_1,\cdots,x_n)'$ (可能是复虚的)，即

$$\sum_{j=1}^{n}a_{ij}x_j=q_1x_i,$$

则

$$|q_1||x_i|\leqslant\sum_{j=1}^{n}a_{ij}|x_j|\leqslant q\max(|x_1|,\cdots,|x_n|),$$

取 $|x_i|=\max(|x_1|,\cdots,|x_n|)$，则得

$$|q_1|\leqslant q.$$

定理证完.

定理 3 对应于高标 q 的特征(列)矢量只有一个(如果不计一个常数因子的话).

证明 若在 $e'=(1,1,\cdots,1)'$ 之外，还有一个 x'，则 x' 是实矢量，对于充分小的 α,

$$e' + \alpha x'$$

仍然是一个非负的，而且对应于特征值 q 的特征矢量．因此 $e' + \alpha x' = \beta e'$，即 x' 是 e' 的常数倍，因而得出本定理．

§5. 标准型方阵的四则运算

定理 1 两个标准型非负方阵之和仍然是标准型，其高标等于两高标之和；其积仍为标准型，其高标等于两高标之积．

证明 如果

$$\sum_{j=1}^{n} a_{ij} = q, \quad \sum_{j=1}^{n} b_{ij} = r,$$

则

$$\sum_{j=1}^{n} (a_{ij} + b_{ij}) = q + r,$$

$$\sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} b_{jk} \right) a_{ij} = qr.$$

如果暂不管非负性，则 A^{-1} 的行和等于 q^{-1}．证明此点是不难的．原因是

$$A(1,1,\cdots,1)' = q(1,1,\cdots,1)',$$

所以

$$A^{-1}(1,1,\cdots,1)' = q^{-1}(1,1,\cdots,1)'.$$

因而 A^{-1} 的行和等于 q^{-1}．

定理 2 如果 $A \geqslant 0$ 及 $(I - A)^{-1} \geqslant 0$，则 A 的高标 <1，而且 $(I - A)^{-1}$ 也与 A 同为标准型，高标是 $(1 - q)^{-1}$．

证明 从

$$A(1,1,\cdots,1)' = q(1,1,\cdots,1)'$$

得

$$(I - A)^{-1}(1,1,\cdots,1)' = (1 - q)^{-1}(1,1,\cdots,1)'.$$

如果 $(1 - A)^{-1} \geqslant 0$，则 $q < 1$．因而得出定理．

同法不难推得，如果 $A \geqslant 0$，其高标等于 q，则 e^A 的高标就是 e^q．而且 A 与 e^A 的正特征矢量相同．

以上所说的虽然是关于列特征矢量的，但同样适用于行特征矢量．

分别命一个非负不可分拆方阵 A 的一个正特征(行)矢量及一个正特征(列)矢量为 x 及 y，即

$$xA = qx, \quad Ay = yq.$$

如此则

$$z_1 = x_1 y_1, \cdots, z_n = x_n y_n.$$

是经过 $\Lambda A \Lambda^{-1}$ 的运算而不变的．

行和相等的标准型，其列特征矢量等于$(1,1,\cdots,1)'$．而行特征矢量是$(z_1,\cdots,z_n)$．同样如果研究列和相等的标准型，则其行特征矢量等于 $(1, 1,\cdots,1)$，而列特征矢量是 $(z_1,\cdots,z_n)'$．

§6. 方阵大小

定理 1 命

$$C=(c_{ij}) \tag{1}$$

是一个复元素方阵，如果

$$|c_{ij}|\leqslant a_{ij}, \tag{2}$$

而 $A=(a_{ij})$ 是不可分拆的非负方阵，其高标等于 q，则 C 的任一特征根 γ 的绝对值都不超过 q，即

$$|\gamma|\leqslant q. \tag{3}$$

如果 $\gamma=e^{i\theta}q$，则

$$C=e^{i\theta}[e^{i\theta_1},\cdots,e^{i\theta_n}]A[e^{i\theta_1},\cdots,e^{i\theta_n}]^{-1}.$$

证明

1) 并不失去普遍性，我们可以假定 A 是标准型. 命 γ 与 $x(\neq 0)$ 是 C 的特征根及其对应的特征(列)矢量，即

$$\gamma x_i=\sum_{j=1}^{n}c_{ij}x_j.$$

由假定可知

$$|\gamma||x_i|\leqslant\sum_{j=1}^{n}|c_{ij}||x_j|\leqslant\sum_{j=1}^{n}a_{ij}|x_j|\leqslant q\max(|x_1|,\cdots,|x_n|), \tag{4}$$

取 $|x_i|=\max(|x_1|,\cdots,|x_n|)$，即得 $|\gamma|\leqslant q$.

2) 假定 $\gamma=e^{i\theta}q$，我们现在检查不等式(4)的各个环节，不失普遍性可以假定

$$|x_1|=|x_2|=\cdots=|x_s|>|x_{s+1}|\geqslant\cdots\geqslant|x_n|.$$

于是，当 $i=1,2,\cdots,s$ 时，(4)式左边等于右边，因此

$$q|x_i|=\sum_{j=1}^{n}|c_{ij}||x_j|=\sum_{j=1}^{n}a_{ij}|x_j|=\left(\sum_{j=1}^{n}a_{ij}\right)|x_i|.$$

最后一个等式仅当

$$a_{is+1}=a_{is+2}=\cdots=a_{in}=0,\qquad 1\leqslant i\leqslant s$$

时才能成立. 也就是 A 是可分拆的，这与假定相违背. 因此

$$|x_1|=|x_2|=\cdots=|x_n|.$$

并且不妨假定 $|x_i|=1$. 因此

$$x_1=e^{i\theta_1},\cdots,x_n=e^{i\theta_n}.$$

方阵

$$C_1=[e^{i\theta_1},\cdots,e^{i\theta_n}]^{-1}\ C[e^{i\theta_1},\cdots,e^{i\theta_n}]$$

是以 $(1,1,\cdots,1)'$ 为其特征(列)矢量的. 我们不妨假定 C_1 就是 C，即假定 C 对应于特征根 $\gamma=qe^{i\theta}$，有一特征(列)矢量 $(1,1,\cdots,1)'$，因此

$$\gamma=qe^{i\theta}=\sum_{j=1}^{n}c_{ij},\quad |c_{ij}|=a_{ij},\quad \sum_{j=1}^{n}a_{ij}=q.$$

即

$$q=\sum_{j=1}^{n}c_{ij}e^{-i\theta}=\sum_{j=1}^{n}a_{ij},$$

由此得出

$$c_{ij}=a_{ij}e^{i\theta}.$$

即得所证.

定理 2 命 A 是一个不可分拆的非负方阵，A 的任一主子方阵的高标一定小于 A 的高标.

证明 命

$$A_1=\begin{pmatrix} a_{11} & \cdots a_{1m} & 0\cdots 0 \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots a_{mm} & 0\cdots 0 \\ 0\cdots & \cdots 0 & 0\cdots 0 \\ \cdots & \cdots & \cdots \end{pmatrix},$$

显然 $A_1\leqslant A$ 而 $A_1\neq A$，由此由定理 1 可知 A_1 的高标小于 A 的高标.

定理 3 命 A 为一不可分拆的非负方阵，则它的高标不是它的特征方程的重根.

证明 特征多项式

$$f(\lambda)=\begin{vmatrix} \lambda-a_{11}, & -a_{12}, & \cdots, & -a_{1n} \\ -a_{21}, & \lambda-a_{22}, & \cdots, & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1}, & -a_{n2}, & \cdots, & \lambda-a_{nn} \end{vmatrix}$$

的微商等于

$$f'(\lambda)=\begin{vmatrix} 1, & 0, & \cdots, & 0 \\ -a_{21}, & \lambda-a_{22}, & \cdots, & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1}, & -a_{n2}, & \cdots, & \lambda-a_{nn} \end{vmatrix}+\begin{vmatrix} \lambda-a_{11}, & -a_{12}, & \cdots, & -a_{1n} \\ 0 & 1, & \cdots, & 0 \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1}, & -a_{n2}, & \cdots, & \lambda-a_{nn} \end{vmatrix}+\cdots,$$

其第一个行列式是主子方阵

$$\begin{pmatrix} a_{22}, & \cdots, & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{n2}, & \cdots, & a_{nn} \end{pmatrix}$$

的特征多项式. 由定理 2 可知它的高标小于 A 的高标 q. 也就当 $\lambda\geqslant q$ 时，它的数值是正,同样处理其它各项. 因而得出: 当 $\lambda\geqslant q$ 时

$$f'(\lambda)>0.$$

也就是 q 是 $f(\lambda)$ 的单根.

定理 4 如果 A 还有绝对值等于 q 的特征根，则它们是

$$qe^{2\pi i\frac{l}{h}},\ l=0,1,2,\cdots,h-1.$$

这儿 h 是一个$\geqslant 2$ 的正整数.

证明 取 $\gamma=qe^{i\theta}$，则由定理 1 的第二部份可知(取 $C=A$)，

$$A=e^{i\theta}[e^{i\theta_1},\cdots,e^{i\theta_n}]A[e^{i\theta_1},\cdots,e^{i\theta_n}]^{-1}. \tag{5}$$

因此得出，如果 x_0 是 A 的特征根，则 $\bar{e}^{i\theta}x_0$ 也是，因而

$$x_0,\ \bar{e}^{i\theta}x_0,\ \bar{e}^{2i\theta}x_0,\cdots$$

都是，但不可能有无穷个特征根，因此有 h 使 $h\theta$ 是 2π 的倍数．命 h 是其中的最小正数，由于无重根的性质，立刻可以推出，A 的所有的绝对值等于 q 的特征根如定理所述．

由于 A^h 出现了重根，因此

定理 5 如果 A 除 q 之外，还有绝对值等于 q 的特征根，则有正整数 h 使 A^h 是可分拆的．

再进一步考虑 (5) 式：

$$A = e^{\frac{2\pi i}{h}}[e^{i\theta_1},\cdots,e^{i\theta_n}]A[e^{i\theta_1},\cdots,e^{i\theta_n}]^{-1}, \tag{6}$$

迭代 h 次，得

$$A = [e^{ih\theta_1},\cdots,e^{ih\theta_n}]A[e^{ih\theta_1},\cdots,e^{ih\theta_n}]^{-1}.$$

如果有一个 j，使 $e^{ih\theta_j} \neq 1$．则 A 可分拆，因此 $e^{i\theta_j}$ 也都是 1 的 h 次方根．假定 $e^{i\theta_1},\cdots,e^{i\theta_n}$ 中有 n_1 个 1，n_2 个 $e^{\frac{2\pi i}{h}}$，n_3 个 $e^{2\cdot\frac{2\pi i}{h}}$ 等等，经排列后假定

$$[e^{i\theta_1},\cdots,e^{i\theta_n}] = \begin{pmatrix} I^{(n_1)} & & & \\ & e^{\frac{2\pi i}{h}}I^{(n_2)} & & 0 \\ & & \ddots & \\ 0 & & & e^{\frac{2\pi i}{h}(h-1)}I^{(n_h)} \end{pmatrix},$$

把 A 拆成为

$$\begin{pmatrix} A_{11}, & A_{12}, & \cdots, & A_{1h} \\ \cdots & \cdots & \cdots & \cdots \\ A_{h1}, & A_{h2}, & \cdots, & A_{hh} \end{pmatrix},$$

代入 (6) 式可知

$$A_{st} = e^{\frac{2\pi i}{h}}\cdot e^{\frac{2\pi i}{h}(s-1)}\cdot e^{-\frac{2\pi i}{h}(t-1)}A_{st} = e^{\frac{2\pi i}{h}(1+s-t)}A_{st}.$$

即当 $t \neq 1+s$ 时，$A_{st} = 0$．即 A 形如

$$\begin{pmatrix} 0, & A_{12}, & 0, & \cdots, & 0 \\ 0, & 0, & A_{23}, & \cdots, & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{h1}, & & 0, & \cdots, & 0 \end{pmatrix}.$$

§7. 强不可拆方阵

定义 1 一个任意幂都不能分拆的方阵称为强不可分拆的．

关于强不可分拆方阵有以下的性质：

定理 1 如果 A 是强不可分拆的非负方阵，而 q 为高标，则

$$\lim_{l\to\infty}\left(\frac{A}{q}\right)^l = u'v,\ vu' = 1. \tag{1}$$

这儿 u' 与 v 分别是 A 的正的列特征矢量与正的行特征矢量．

证明 由于强不可分拆，在 A 中 q 以外的其他特征根的绝对值都 $< q$．因此 A/q 有一个特征根等于 1，而其它的都<1．所以

$$\lim_{l\to\infty}\left(\frac{A}{q}\right)^l$$

是一个方阵，有一个特征根等于 1（而且是单根），其它的等于 0，因此

$$\lim_{l\to\infty}\left(\frac{A}{q}\right)^l = u'v, \quad vu' = 1.$$

命 c 是任一矢量，命

$$v_l = c\frac{A^l}{q^l}, \tag{2}$$

则

$$\lim_{l\to\infty} v_l = (cu')v. \tag{3}$$

即不管 c 如何，v_l 的极限是 v 乘上一个常数因子，又从

$$qv_{l+1} = c\frac{A^{l+1}}{q^l} = v_l A$$

及 $l\to\infty$ 可知

$$qv = vA.$$

即 v 是行特征矢量.

同法证明

$$qu' = Au',$$

因此 u, v 都是正矢量.

由此立刻推出:

定理 2 对任一强不可分拆方阵 A，有一正整数 l 存在，使

$$A^l > 0.$$

即 A^l 的每个元数都是正的.

§ 8. Марков 链

定义 1 一个适合于

$$x_1 + \cdots + x_n = 1, \quad x_i \geqslant 0$$

的矢量 $x = (x_1, \cdots, x_n)$ 称为概率矢量，一个方阵

$$P = (p_{ij})_{1\leqslant i,j\leqslant n}, \quad p_{ij} \geqslant 0, \quad \sum_{i=1}^{n} p_{ij} = 1,$$

称为 Марков 方阵.

Марков 方阵把概率列矢量变为概率列矢量，即 Px' 仍然是概率列矢量. 理由是由

$$y_i = \sum_{j=1}^{n} p_{ij}x_j,$$

可知

$$\sum_{i=1}^{n} y_i = \sum_{j=1}^{n}\sum_{i=1}^{n} p_{ij}x_j = \sum_{j=1}^{n} x_j = 1.$$

我们现在研究一个物理系统，这个系统只可能有有限种“态”，而且只可能在某些一定的时间变态.

我们用 $1, 2, \cdots, n$ 来表示这些不同的态，而当 $t = 0, 1, 2, \cdots$ 时变“态”在时间 t 这

个物理系统是态 j，而到时间 $t+1$，这个物理系统变为态 i 的可能性以概率

$$p_{ij}, \quad \sum_{j=1}^{n} p_{ij} = 1$$

表之.

命 $x_i(t)$ 表示在时间 t，第 i 态出现的概率，如此则有

$$x_i(t+1) = \sum_{j=1}^{n} p_{ij} x_j(t).$$

写成方阵符号则

$$x(t+1)' = Px(t)'.$$

并假定初始态是

$$x(0) = (c_1, \cdots, c_n), \quad c_1 + \cdots + c_n = 1.$$

Марков 的基本定理是:

定理 1 如果 $P > 0$，则

$$\lim_{t\to\infty} x(t) = y.$$

y 仍然是一个概率矢量，它是 P 的唯一正特征矢量，而与初始态 $x(0)$ 无关.

证明 这是定理 7.1 的显然推理，由于 P 是正的，因此是强不可分拆的. 所以

$$\lim_{l\to\infty} P^l = u'v, \; vu' = 1. \tag{1}$$

由于 P 是 Марков 方阵，以 $(1,1,\cdots,1)$ 为其(行)特征矢量. 因此，$v = (1,1,\cdots,1)$.

$$\lim_{t\to\infty} x(t)' = \lim_{t\to\infty} P^t x(0)' = u'vx(0)' = u'.$$

即得

$$\lim_{t\to\infty} x(t) = u.$$

以上的定理不仅对正方阵 P 正确，显然对强不可分拆的方阵也是正确的. 关于不可分拆方阵，由于 P 可能有绝对值等于 1 的特征根 $e^{2\pi jk/h}$，$k = 0,1,2,\cdots,h-1$. 因而不能得出结论(1)来. 虽然如此，我们考虑算术平均

$$\frac{1}{L}\sum_{l=1}^{L} A^l. \tag{2}$$

由于

$$\lim_{L\to\infty} \frac{1}{L}\sum_{l=1}^{L} e^{2\pi ik\frac{l}{h}} = \begin{cases} 0, \text{如果 } h \text{ 除不尽 } k; \\ 1, \text{如果 } h \text{ 除得尽 } k. \end{cases}$$

因此，当 $L\to\infty$，(2) 也是一个秩等于 1，而且有 1 为特征根的方阵. 因此

$$\lim_{L\to\infty} \frac{1}{L}\sum_{l=1}^{L} P^l = u'v, \quad vu' = 1.$$

由此立得

$$\lim_{L\to\infty} \frac{1}{L}\sum_{l=1}^{L} vP^l = v.$$

因而 $vP = v$，即 $v = (1,1,\cdots,1)$. 所以

$$\lim_{L\to\infty} \frac{1}{L}\sum_{l=1}^{L} P^l x(0)' = u'(vx(0)') = u'.$$

即

$$\lim_{L\to\infty}\frac{1}{L}\sum_{l=1}^{L}x(l)=u.$$

定理 2 如果 P 是不可分拆的 Марков 方阵，则

$$\lim_{L\to\infty}\frac{1}{L}\sum_{l=1}^{L}x(l)=y.$$

y 仍然是一个概率矢量，它是 P 的唯一的正特征矢量，而且与初始态 $x(0)$ 无关。

§9. 连续随机过程

我们现在把上节的工作从离散的情况转化为连续变化的情况，把时间分为

$$t=0,\Delta,2\Delta,\cdots.$$

为了要使连续的过程有意义，我们假定当时间愈小时，愈接近于不变；也就是我们假定（当△是充分小时）：$a_{ij}\Delta$ 是系统 S 由时间 t 的 j 变态为时间 $t+\Delta$ 时的 i 态的概率 $(i\neq j)$ 及 $1-a_{jj}\Delta$ 是系统 S 中由时间 t 的 j 态变为时间 $t+\Delta$ 的 j 态的概率。

此处

$$a_{ij}\geqslant 0,\quad i\neq j$$

$$a_{jj}=\sum_{\substack{i=1\\ i\neq j}}^{n}a_{ij}. \tag{1}$$

这样便得

$$x_i(t+\Delta)=(1-a_{ii}\Delta)x_i(t)+\Delta\sum_{j\neq i}a_{ij}x_j(t),\quad i=1,2,\cdots,n$$

及 $t=0,\Delta,2\Delta,\cdots$.

假定

$$x_i(t+\Delta)=x_i(t)+\Delta\frac{dx_i(t)}{dt}+O(\Delta^2),$$

则当 $\Delta\to 0$ 时有

$$\frac{dx_i}{dt}=-a_{ii}x_i+\sum_{\substack{i=1\\ j\neq i}}^{n}a_{ij}x_j,\ x_i(0)=c_i,\quad i=1,2,\cdots,n. \tag{2}$$

此处 $C=(c_1,\cdots,c_n)$ 是概率矢量

$$c_i\geqslant 0\quad \sum_{i=1}^{n}c_i=1.$$

这个微分方程称为过程微分方程．命

$$M=\begin{pmatrix}-a_{11}, & a_{12}, & \cdots, & a_{1n}\\ a_{21}, & -a_{22}, & \cdots, & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1}, & a_{n2}, & \cdots, & -a_{nn}\end{pmatrix},\quad a_{jj}=\sum_{\substack{i=1\\ i\neq j}}^{n}a_{ij},\ a_{ij}\geqslant 0\quad (i\neq j). \tag{3}$$

则方程 (2) 可以写成为

$$\frac{dx}{dt} = xM', \quad x(0) = c. \tag{4}$$

它的解答显然是

$$x = ce^{M't}. \tag{5}$$

我们现在证明，对任一初始概率矢量 c，x 也一定是概率矢量，要证这点十分容易. 命 $u = (1, 1, \cdots, 1)$. 由(3)可见

$$uM = 0.$$

因此

$$ue^{Mt} = u\left(I + Mt + \frac{M^2}{2!}t^2 + \cdots\right) = u$$

及

$$xu' = ce^{M't}u' = cu' = 1.$$

又由第三章 § 7 的定理 3 可知 $e^{M't}$ 是非负元素方阵. 因而得出

定理 1 过程微分方程的解是概率矢量.

现在考虑 $t \to \infty$ 时过程方程的解的性质.

定理 2 如果 $a_{ij} > 0$ $(i \neq j)$，则当 $t \to \infty$ 时，过程方程的解所趋近的矢量与初始矢量无关.

证明 不难证明在新添的条件 $(a_{ij} > 0\ (i \neq j))$ 下，$e^M > 0$. 因此 e^M 只有一个特征根等于 1，其他的绝对值都小于 1，因此

$$\lim_{t \to \infty} (e^M)^t = p'q, \quad pq' = 1.$$

由 $ue^{Mt} = u$，可知 $up'q = u$，即 $(up')q = u$. 命 $\frac{1}{up'}p = v$，则得

$$\lim_{t \to \infty} (e^M)^t = v'u, \quad vu' = 1.$$

由此推出

$$\lim_{t \to \infty} x(t) = \lim_{t \to \infty} ce^{M't} = \lim_{t \to \infty} cu'v = v.$$

即得所证.

思 考 问 题

1. 定理 2 的条件能否减弱，例如：我们引进 M 类型方阵的等价关系(即在广义换置方阵下相似)，及不可分拆的概念，是否对不可分拆的 M，定理 2 仍然正确.

2. 概率方阵 P 与适合于 (3) 的方阵 M 之间能否建立关系 $P = e^M$ 如果不能试找出例外，并找出可能的条件来，如果 M 可分拆显然 P 可分拆，逆定理正确否?

［Ｇｅｎｅｒａｌ　Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝高等数学引论　余篇
作者＝华罗庚
页数＝１７１
ＳＳ号＝１０１７９７９７
出版日期＝１９８４年０７月第１版